



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 9/60		(11) International Publication Number: WO 00/60059
		(43) International Publication Date: 12 October 2000 (12.10.00)
<p>(21) International Application Number: PCT/DK00/00148</p> <p>(22) International Filing Date: 26 March 2000 (26.03.00)</p> <p>(30) Priority Data: PA 1999 00437 30 March 1999 (30.03.99) DK</p> <p>(71) Applicant: NOVO NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsværd (DK).</p> <p>(72) Inventors: ANDERSEN, Carsten; Højeloft Vænge 162, DK-3500 Værløse (DK). JØRGENSEN, Christel, Thea; Livæggergade 41, 4th, DK-2100 Copenhagen Ø (DK). BISGARD-FRANTZEN, Henrik; Elmevængts 8B, DK-2880 Bagsværd (DK). SVENDSEN, Allan; Bakkedædet 28, DK-3460 Birkerød (DK). KJÆRULFF, Soren; Kongsdalsvej 47, DK-2720 Værløse (DK).</p>		<p>(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIGO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TI, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BP, BI, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p>Published <i>Without international search report and to be republished upon receipt of that report.</i></p>
<p>(54) Title: ALPHA-AMYLASE VARIANTS</p> <p>(57) Abstract</p> <p>The invention relates to a variant of a parent Thermamyl-like alpha-amylase, which variant exhibits altered properties, in particular reduced capability of cleaving a substrate close to the branching point, and improved substrate specificity and/or improved specific activity relative to the parent alpha-amylase.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lithuania	SI	Slovenia
AM	Amenica	FJ	Fiji	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LG	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TB	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MH	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Malta	TT	Trinidad and Tobago
BJ	Burkina	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BX	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LJ	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LX	Liberia	SG	Singapore		

Alpha-AMYLASE VARIANTS

FIELD OF THE INVENTION

The present invention relates, inter alia, to novel variants of parent Termamyl-like alpha-amylases, notably variants exhibiting altered properties, in particular altered cleavage pattern (relative to the parent) which are advantageous with respect to applications of the variants in, in particular, industrial starch processing (e.g., starch liquefaction or saccharification).

BACKGROUND OF THE INVENTION

Alpha-Amylases (alpha-1,4-glucan-4-glucanhydrolases, EC 3.2.1.1) constitute a group of enzymes which catalyze hydrolysis of starch and other linear and branched 1,4-glucosidic oligo- and polysaccharides.

There is a very extensive body of patent and scientific literature relating to this industrially very important class of enzymes. A number of alpha-amylase such as Termamyl-like alpha-amylases variants are known from, e.g., WO 90/11352, WO 95/10603, WO 95/26397, WO 96/23873, WO 96/23874 and WO 97/41213.

Among recent disclosure relating to alpha-amylases, WO 96/23874 provides three-dimensional, X-ray crystal structural data for a Termamyl-like alpha-amylase, referred to as BA2, which consists of the 300 N-terminal amino acid residues of the *B. amyloliquefaciens* alpha-amylase comprising the amino acid sequence shown in SEQ ID NO: 6 herein and amino acids 301-483 of the C-terminal end of the *B. licheniformis* alpha-amylase comprising the amino acid sequence shown in SEQ ID NO: 4 herein (the latter being available commercially under the trademark Termamyl™), and which is thus closely related to the industrially important *Bacillus* alpha-amylases (which in the present context are embraced within the meaning of the term "Termamyl-like alpha-amylases", and which include, inter alia, the *B. licheniformis*, *B. amyloliquefaciens* and *B. stearothermophilus* alpha-amylases). WO 96/23874 further describes methodology for designing, on the basis of an analysis

of the structure of a parent Termamyl-like alpha-amylase, variants of the parent Termamyl-like alpha-amylase which exhibit altered properties relative to the parent.

WO 96/23874 and WO 97/41213 (Novo Nordisk) discloses Termamyl-like alpha-amylase variants with an altered cleavage pattern containing mutations in the amino acid residues V54, D53, Y56, Q333, G57 and A52 of the sequence shown in SEQ ID NO: 4 herein.

10 BRIEF DISCLOSURE OF THE INVENTION

The present invention relates to novel alpha-amylolytic variants (mutants) of a Termamyl-like alpha-amylase, in particular variants exhibiting altered cleavage pattern (relative to the parent), which are advantageous in connection with the industrial processing of starch (starch liquefaction, saccharification and the like).

The inventors have surprisingly found variants with altered properties, in particular altered cleavage pattern which have improved reduced capability of cleaving an substrate close to the branching point, and further have improved substrate specificity and/or improved specific activity, in comparison to the WO 96/23874 and WO 97/41213 (Novo Nordisk) disclosed Termamyl-like alpha-amylase variants with an altered cleavage pattern containing mutations in the amino acid residues V54, D53, Y56, Q333, G57 and A52 of the sequence shown in SEQ ID NO: 4 herein.

The invention further relates to DNA constructs encoding variants of the invention, to composition comprising variants of the invention, to methods for preparing variants of the invention, and to the use of variants and compositions of the invention, alone or in combination with other alpha-amylolytic enzymes, in various industrial processes, e.g., starch liquefaction, and in detergent compositions, such as laundry, dish washing and hard surface cleaning compositions; ethanol production, such as fuel, drinking and industrial ethanol production; desizing of textiles, fabrics or garments etc.

Nomenclature

In the present description and claims, the conventional one-letter and three-letter codes for amino acid residues are used.

For ease of reference, alpha-amylase variants of the invention are described by use of the following nomenclature:

Original amino acid(s) :position(s) :substituted amino acid(s)

According to this nomenclature, for instance the substitution of alanine for asparagine in position 30 is shown as:

18 Ala30Asn or A30N

a deletion of alanine in the same position is shown as:

Ala30* or A30*

and insertion of an additional amino acid residue, such as lysine, is shown as:

19 *30alys or *30aL

A deletion of a consecutive stretch of amino acid residues, such as amino acid residues 30-33, is indicated as (30-33)* or Δ(A30-N33) or delta(A30-N33).

Where a specific alpha-amylase contains a "deletion" in comparison with other alpha-amylases and an insertion is made in such a position this is indicated as:

*36aAsp or *36aD

for insertion of an aspartic acid in position 36

Multiple mutations are separated by plus signs, i.e.:

20 Ala30Asp + Glu34Ser or A30N+E34S

representing mutations in positions 30 and 34 substituting alanine and glutamic acid for asparagine and serine, respectively. Multiple mutations may also be separated as follows, i.e., meaning the same as the plus sign:

21 Ala30Asp/Glu34Ser or A30N/E34S

When one or more alternative amino acid residues may be inserted in a given position it is indicated as

A30N, E or

A30N or A30E

22 Furthermore, when a position suitable for modification is identified herein without any specific modification being suggested, or A30X, it is to be understood that any amino acid

residue may be substituted for the amino acid residue present in the position. Thus, for instance, when a modification of an alanine in position 30 is mentioned, but not specified, or specified as "X", it is to be understood that the alanine may be deleted or substituted for any other amino acid, i.e., any one of: R,N,D,C,Q,E,G,H,I,L,K,M,F,P,S,T,W,Y,V.

DETAILED DISCLOSURE OF THE INVENTION

The Termamyl-like alpha-amylase

It is well known that a number of alpha-amylases produced by *Bacillus* spp. are highly homologous on the amino acid level. For instance, the *B. licheniformis* alpha-amylase comprising the amino acid sequence shown in SEQ ID NO: 4 (commercially available as Termamyl™) has been found to be about 89% homologous with the *B. amyloliquefaciens* alpha-amylase comprising the amino acid sequence shown in SEQ ID NO: 6 and about 79% homologous with the *B. stearothermophilus* alpha-amylase comprising the amino acid sequence shown in SEQ ID NO: 8. Further homologous alpha-amylases include an alpha-amylase derived from a strain of the *Bacillus* sp. NCIB 12289, NCIB 12512, NCIB 12513 or DSM 9375, all of which are described in detail in WO 95/26397, and the #707 alpha-amylase described by Tsukamoto et al., Biochemical and Biophysical Research Communications, 151 (1988), pp. 25-31.

Still further homologous alpha-amylases include the alpha-amylase produced by the *B. licheniformis* strain described in EP 0252666 (ATCC 27811), and the alpha-amylases identified in WO 91/00353 and WO 94/18314. Other commercial Termamyl-like *B. licheniformis* alpha-amylases are Optitherm™ and Takatherm™ (available from Solvay), Maxamyl™ (available from Gist-brocades/Genencor), Spezyme AA™ and Spezyme Delta AA™ (available from Genencor), and Keistase™ (available from Daiwa).

Because of the substantial homology found between these alpha-amylases, they are considered to belong to the same class of alpha-amylases, namely the class of "Termamyl-like alpha-amylases".

Accordingly, in the present context, the term "Termamyl-like alpha-amylase" is intended to indicate an alpha-amylase, which at the amino acid level exhibits a substantial homology to Termamyl™, i.e., the *B. licheniformis* alpha-amylase having the amino acid sequence shown in SEQ ID NO: 4 herein. In other words, a Termamyl-like alpha-amylase is an alpha-amylase, which has the amino acid sequence shown in SEQ ID NO: 2, 4, 6, or 8 herein, and the amino acid sequence shown in SEQ ID NO: 1 or 2 of WO 95/26397 or in Tsukamoto et al., 1983, or i) which displays at least 60%, preferred at least 70%, more preferred at least 75%, even more preferred at least 80%, especially at least 85%, especially preferred at least 90%, even especially more preferred at least 95% homology, more preferred at least 97%, more preferred at least 99% with at least one of said amino acid sequences and/or ii) displays immunological cross-reactivity with an antibody raised against at least one of said alpha-amylases, and/or iii) is encoded by a DNA sequence which hybridises to the DNA sequences encoding the above-specified alpha-amylases which are apparent from SEQ ID NOS: 1, 3, 5 and 7 of the present application and SEQ ID NOS: 4 and 5 of WO 95/26397, respectively.

In connection with property i), the "homology" may be determined by use of any conventional algorithm, preferably by use of the GAP programme from the GCG package version 7.3 (June 1993) using default values for GAP penalties, which is a GAP creation penalty of 3.0 and GAP extension penalty of 0.1, (Genetic Computer Group (1991) Programme Manual for the GCG Package, version 7, 575 Science Drive, Madison, Wisconsin, USA 53711).

A structural alignment between Termamyl and a Termamyl-like alpha-amylase may be used to identify equivalent/corresponding positions in other Termamyl-like alpha-amylases. One method of obtaining said structural alignment is to use the Pile Up programme from the GCG package using default values of gap penalties, i.e., a gap creation penalty of 3.0 and gap extension penalty of 0.1. Other structural alignment methods include the hydrophobic cluster analysis (Gaboriaud et al.,

(1987), FEBS LETTERS 224, pp. 149-155) and reverse threading (Huber, T ; Torda, AE, PROTEIN SCIENCE Vol. 7, No. 1 pp. 142-149 (1998). Property ii) of the alpha-amylase, i.e., the immunological cross reactivity, may be assayed using an antibody raised against, or reactive with, at least one epitope of the relevant Termamyl-like alpha-amylase. The antibody, which may either be monoclonal or polyclonal, may be produced by methods known in the art, e.g., as described by Hudson et al., Practical Immunology, Third edition (1989), Blackwell Scientific Publications. The immunological cross-reactivity may be determined using assays known in the art, examples of which are Western Blotting or radial immunodiffusion assay, e.g., as described by Hudson et al., 1989. In this respect, immunological cross-reactivity between the alpha-amylases having the amino acid sequences SEQ ID NOS: 2, 4, 6, or 8, respectively, have been found.

The oligonucleotide probe used in the characterization of the Termamyl-like alpha-amylase in accordance with property iii) above may suitably be prepared on the basis of the full or partial nucleotide or amino acid sequence of the alpha-amylase in question.

Suitable conditions for testing hybridization involve presoaking in 6xSSC and prehybridizing for 1 hour at ~40°C in a solution of 20% formamide, 5xDenhardt's solution, 50mM sodium phosphate, pH 6.8, and 50mg of denatured sonicated calf thymus DNA, followed by hybridization in the same solution supplemented with 100mM ATP for 18 hours at ~40°C, followed by three times washing of the filter in 2xSSC, 0.2% SDS at 40°C for 30 minutes (low stringency), preferred at 50°C (medium stringency), more preferably at 65°C (high stringency), even more preferably at ~75°C (very high stringency). More details about the hybridization method can be found in Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, 1989.

In the present context, "derived from" is intended not only to indicate an alpha-amylase produced or producible by a strain of the organism in question, but also an alpha-amylase encoded

by a DNA sequence isolated from such strain and produced in a host organism transformed with said DNA sequence. Finally, the term is intended to indicate an alpha-amylase, which is encoded by a DNA sequence of synthetic and/or cDNA origin and which has the identifying characteristics of the alpha-amylase in question. The term is also intended to indicate that the parent alpha-amylase may be a variant of a naturally occurring alpha-amylase, i.e., a variant, which is the result of a modification (insertion, substitution, deletion) of one or more amino acid residues of the naturally occurring alpha-amylase.

Parent hybrid alpha-amylases

The parent alpha-amylase may be a hybrid alpha-amylase, i.e., an alpha-amylase, which comprises a combination of partial amino acid sequences derived from at least two alpha-amylases.

The parent hybrid alpha-amylase may be one, which on the basis of amino acid homology and/or immunological cross-reactivity and/or DNA hybridization (as defined above) can be determined to belong to the Termamyl-like alpha-amylase family. In this case, the hybrid alpha-amylase is typically composed of at least one part of a Termamyl-like alpha-amylase and part(s) of one or more other alpha-amylases selected from Termamyl-like alpha-amylases or non-Termamyl-like alpha-amylases of microbial (bacterial or fungal) and/or mammalian origin.

Thus, the parent hybrid alpha-amylase may comprise a combination of partial amino acid sequences deriving from at least two Termamyl-like alpha-amylases, or from at least one Termamyl-like and at least one non-Termamyl-like bacterial alpha-amylase, or from at least one Termamyl-like and at least one fungal alpha-amylase. The Termamyl-like alpha-amylase from which a partial amino acid sequence derives may, e.g., be any of those specific Termamyl-like alpha-amylases referred to herein.

For instance, the parent alpha-amylase may comprise a C-terminal part of an alpha-amylase derived from a strain of *B. licheniformis*, and a N-terminal part of an alpha-amylase derived from a strain of *B. amyloliquefaciens* or from a strain of *B. stearothermophilus*. For instance, the parent alpha-amylase may

comprise at least 430 amino acid residues of the C-terminal part of the *B. licheniformis* alpha-amylase, and may, e.g., comprise a) an amino acid segment corresponding to the 37 N-terminal amino acid residues of the *B. amyloliquefaciens* alpha-amylase having the amino acid sequence shown in SEQ ID NO: 6 and an amino acid segment corresponding to the 445 C-terminal amino acid residues of the *B. licheniformis* alpha-amylase having the amino acid sequence shown in SEQ ID NO: 4, or b) an amino acid segment corresponding to the 68 N-terminal amino acid residues of the *B. stearothermophilus* alpha-amylase having the amino acid sequence shown in SEQ ID NO: 8 and an amino acid segment corresponding to the 415 C-terminal amino acid residues of the *B. licheniformis* alpha-amylase having the amino acid sequence shown in SEQ ID NO: 4.

In a preferred embodiment the parent Termamyl-like alpha-amylase is a hybrid Termamyl-like alpha-amylase identical to the *Bacillus licheniformis* alpha-amylase shown in SEQ ID NO: 4, except that the N-terminal 35 amino acid residues (of the mature protein) is replaced with the N-terminal 33 amino acid residues of the mature protein of the *Bacillus amyloliquefaciens* alpha-amylase (BAN) shown in SEQ ID NO: 6. Said hybrid may further have the following mutations: N156Y+A181T+N190F+A209V+Q264S (using the numbering in SEQ ID NO: 4) referred to as LE174.

Another preferred parent hybrid alpha-amylase is LE429 shown in SEQ ID NO: 2.

The non-Termamyl-like alpha-amylase may, e.g., be a fungal alpha-amylase, a mammalian or a plant alpha-amylase or a bacterial alpha-amylase (different from a Termamyl-like alpha-amylase). Specific examples of such alpha-amylases include the *Aspergillus oryzae* TAKA alpha-amylase, the *A. niger* acid alpha-amylase, the *Bacillus subtilis* alpha-amylase, the porcine pancreatic alpha-amylase and a barley alpha-amylase. All of these alpha-amylases have elucidated structures, which are markedly different from the structure of a typical Termamyl-like alpha-amylase as referred to herein.

The fungal alpha-amylases mentioned above, i.e., derived from *A. niger* and *A. oryzae*, are highly homologous on the amino

acid level and generally considered to belong to the same family of alpha-amylases. The fungal alpha-amylase derived from *Aspergillus oryzae* is commercially available under the tradename Fungamyl™.

Furthermore, when a particular variant of a Termamyl-like alpha-amylase (variant of the invention) is referred to - in a conventional manner - by reference to modification (e.g., deletion or substitution) of specific amino acid residues in the amino acid sequence of a specific Termamyl-like alpha-amylase, it is to be understood that variants of another Termamyl-like alpha-amylase modified in the equivalent position(s) (as determined from the best possible amino acid sequence alignment between the respective amino acid sequences) are encompassed thereby.

A preferred embodiment of a variant of the invention is one derived from a *B. licheniformis* alpha-amylase (as parent Termamyl-like alpha-amylase), e.g., one of those referred to above, such as the *B. licheniformis* alpha-amylase having the amino acid sequence shown in SEQ ID NO: 4.

Construction of variants of the invention

The construction of the variant of interest may be accomplished by cultivating a microorganism comprising a DNA sequence encoding the variant under conditions which are conducive for producing the variant. The variant may then subsequently be recovered from the resulting culture broth. This is described in detail further below.

Altered properties

The following discusses the relationship between mutations, which may be present in variants of the invention, and desirable alterations in properties (relative to those of a parent Termamyl-like alpha-amylase), which may result therefrom.

In the first aspect the invention relates to a variant of a parent Termamyl-like alpha-amylase, comprising an alteration at one or more positions selected from the group of:

W13, G48, T49, S50, Q51, A52, D53, V54, G57, C107, C108, A111,

S168, M197, wherein (a) the alteration(s) are independently
(i) an insertion of an amino acid downstream of the amino acid which occupies the position,
s (ii) a deletion of the amino acid which occupies the position, or
s (iii) a substitution of the amino acid which occupies the position with a different amino acid,
10 (b) the variant has alpha-amylase activity and (c) each position corresponds to a position of the amino acid sequence of the parent Termamyl-like alpha-amylase having the amino acid sequence of SEQ ID NO: 4.

In a preferred embodiment the above variants of the invention comprise a mutation in a position corresponding to at least one of the following mutations in the amino acid sequence shown in SEQ ID NO: 4:

V54N, A52S, A52S+V54N, T49L, T49+G107A, A52S+V54N+T49L+G107A, A52S+V54N+T49L, G107A, Q51R, Q51R+A52S, A52N; or
T49F+G107A, T49V+G107A, T49D+G107A, T49Y+G107A, T49S+G107A, T49N+G107A, T49I+G107A, T49L+A52S+G107A, T49L+A52T+G107A, 20 T49L+A52F+G107A, T49L+A52L+G107A, T49L+A52I+G107A, T49L+A52V+G107A; or
T49V, T49I, T49D, T49N, T49S, T49Y, T49F, T49W, T49M, T49E, T49Q, T49K, T49R, A52T, A52L, A52I, A52V, A52M, A52F, A52Y, A52W, V54M, G107V, G07I, G107L, G107C.

In a preferred embodiment a variant of the invention comprises at least one mutation in a position corresponding to the following mutations in the amino acid sequence shown in SEQ ID NO: 4:

W13F, L, I, V, Y, A;

G46A, V, S, T, I, L;

*48aD or *48aY (i.e., insertion of D or Y);

T49X;

*49aX (i.e., insertion of any possible amino acid residue)

S50X, in particular D, Y, L, T, V, I;

Q51R, X;

A52X, in particular A52S, N, T, F, L, I, V;

D53E, Q, Y, I, N, S, T, V, L;

V54X, in particular V54I,N,W,Y,F,L;
G57S,A,V,L,I,F,Y,T;
G107X, in particular G107A,V,S,T,I,L,C;
G108X, in particular G108A,V,S,T,I,L;
5 A111V,I,L;
S168Y;
M197X, in particular Y,F,L,I,T,A,G.

In a preferred embodiment a variant of the invention comprises the following mutations corresponding to the following 10 mutations in the amino acid sequence shown in SEQ ID NO: 4:
T49X+A52X+V54N/I/L/Y/F/W+G107A, and may further comprise G108A.

In a preferred embodiment a variant of the invention comprises at least one mutation corresponding to the following mutations in the amino acid sequence shown in SEQ ID NO: 4:

15 T49L+G107A;
T49I+G107A;
T49L+G107A+V54I;
T49I+G107A+V54I;
A52S+V54N+T49L+G107A;
20 A52S+V54I+T49L+G107A;
A52S+T49L+G107A;
A52T+T49L+G107A;
A52S+V54N+T49I+G107A;
A52S+V54I+T49I+G107A;
25 A52S+T49I+G107A;
T49L+G108A;
T49I+G108A;
T49L+G108A+V54I;
T49I+G108A+V54I.

30 All of the above-mentioned variants of the invention have altered properties (meaning increased or decreased properties), in particular at least one of the following properties relative to the parent alpha-amylase: reduced ability to cleave a substrate close to the branching point, improved substrate specificity and/or improved specific activity, altered substrate binding, altered thermal stability, altered pH/activity profile, altered pH/stability profile, altered stability towards

oxidation, altered Ca^{2+} dependency.

5 Stability

In the context of the present invention, mutations (including amino acid substitutions and/or deletions) of importance with respect to achieving altered stability, in particular improved stability (i.e., higher or lower), at especially low pH (i.e., pH 4-6) include any of the mutations listed in the in "Altered properties" section, above and the variants mentioned right below.

The following variants: Q360A,K; N102A, N326A,L, N190G, N190K; Y262A,K,E (using the BAN, i.e., SEQ ID N: 6, numbering) were also tested for pH stability. A preferred parent alpha-amylase may be BA2 described above. The pH stability was determined as described in the "Materials & Methods" section.

20 Ca^{2+} stability

Altered Ca^{2+} stability means the stability of the enzyme under Ca^{2+} depletion has been improved, i.e., higher or lower stability. In the context of the present invention, mutations (including amino acid substitutions) of importance with respect to achieving altered Ca^{2+} stability, in particular improved Ca^{2+} stability, i.e., higher or lower stability, at especially low pH (i.e., pH 4-6) include any of the mutations listed in the in "Altered properties" section above.

25 Specific activity

In a further aspect of the present invention, important mutations with respect to obtaining variants exhibiting altered specific activity, in particular increased or decreased specific activity, especially at temperatures from 60-100°C, preferably 70-95°C, especially 80-90°C, include any of the mutations listed in the in "Altered properties" section above.

The specific activity of LZ174 and LZ429 was determined to

16,000 NU/mg using the Phadebas® assay described in the "Materials and Methods" section.

Altered cleavage pattern

In the starch liquefaction process it is desirable to use an alpha-amylase, which is capable of degrading the starch molecules into long, branched oligosaccharides, rather than an alpha-amylase, which gives rise to formation of shorter, branched oligosaccharides (like conventional Termamyl-like alpha-amylases). Short, branched oligosaccharides (panose precursors) are not hydrolyzed satisfactorily by pullulanases, which are used after alpha-amylase treatment in the liquefaction process, or simultaneously with a saccharifying amyloglucosidase (glucoamylase), or before adding a saccharifying amyloglucosidase (glucoamylase). Thus, in the presence of panose precursors, the product mixture present after the glucoamylase treatment contains a significant proportion of short, branched, so-called limit-dextrin, viz. the trisaccharide panose. The presence of panose lowers the saccharification yield significantly and is thus undesirable.

It has been reported previously (US patent 5,234,623) that, when saccharifying with glucoamylase and pullulanase, the presence of residual alpha-amylase activity arising from the liquefaction process, can lead to lower yields of glucose, if the alpha-amylase is not inactivated before the saccharification stage. This inactivation can be typically carried out by adjusting the pH to below 4.7 at 95°C, before lowering the temperature to 60°C for saccharification.

The reason for this negative effect on glucose yield is not fully understood, but it is assumed that the liquefying alpha-amylase (for example Termamyl 120 L from *B. licheniformis*) generates "limit dextrans" (which are poor substrates for pullulanase), by hydrolysing 1,4-alpha-glucosidic linkages close to and on both sides of the branching points in amylopectin. Hydrolysis of these limit dextrans by glucoamylase leads to a build up of the trisaccharide panose, which is only slowly hydrolysed by

glucosamylase.

The development of a thermostable alpha-amylase, which does not suffer from this disadvantage, would be a significant improvement, as no separate inactivation step would be required.

Thus, the aim of the present invention is to arrive at a mutant alpha-amylase having appropriately modified starch-degradation characteristics but retaining the thermostability of the parent Termamyl-like alpha-amylase.

Accordingly, the invention relates to a variant of a Termamyl-like alpha-amylase, which has an improved reduced ability to cleave a substrate close to the branching point, and further has improved substrate specificity and/or improved specific activity.

Of particular interest is a variant, which cleaves an amylopectin substrate, from the reducing end, more than one glucose unit from the branching point, preferably more than two or three glucose units from the branching point, i.e., at a further distance from the branching point than that obtained by use of a wild type *B. licheniformis* alpha-amylase.

It may be mentioned here that according to WO 96/23874, variants comprising at least one of the following mutations are expected to prevent cleavage close to the branching point:

V54L,I,F,Y,W,R,K,H,E,Q;

D53L,I,F,Y,W;

Y56W;

Q333W;

G57, all possible amino acid residues;

A82, amino acid residues larger than A, e.g., A52W,Y,L,F,I.

Mutations of particular interest in relation to obtaining variants according to the invention having an improved reduced ability to cleave a substrate close to the branching point, and further has improved substrate specificity and/or improved specific activity include mutations at the following positions in *B. licheniformis* alpha-amylase, SEQ ID NO: 4:

H156, A181, N190, A209, Q264 and I201.

It should be emphasized that not only the Termamyl-like

alpha-amylases mentioned specifically below may be used. Also other commercial Termamyl-like alpha-amylases can be used. An unexhaustive list of such alpha-amylases is the following:
Alpha-amylases produced by the *B. licheniformis* strain described in EP 0252666 (ATCC 27811), and the alpha-amylases identified in WO 91/06353 and WO 94/18314. Other commercial Termamyl-like *B. licheniformis* alpha-amylases are Optitherm™ and Takatherm™ (available from Solvay), Maxamyl™ (available from Gist-brocades/Genencor), Spezym AA™ Spezyme Delta AA™ (available from Genencor), and Keistase™ (available from Daiwa).
10

All Termamyl-like alpha-amylase may suitably be used as backbone for preparing variants of the invention.

In a preferred embodiment of the invention the parent Termamyl-like alpha-amylase is a hybrid alpha-amylase of SEQ ID NO: 4 and SEQ ID NO: 6. Specifically, the parent hybrid Termamyl-like alpha-amylase may be a hybrid alpha-amylase comprising the 445 C-terminal amino acid residues of the *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4 and the 37 N-terminal amino acid residues of the mature alpha-amylase derived from *B. amyloliquefaciens* shown in SEQ ID NO: 6, which may suitably further have the following mutations: H156Y+A181T+N190F+A209V+Q264S (using the numbering in SEQ ID NO: 4). This hybrid is referred to as LE174. The LE174 hybrid may be combined with a further mutation I201F to form a parent hybrid Termamyl-like alpha-amylase having the following mutations H156Y+A181T+N190F+A209V+Q264S+I201F (using SEQ ID NO: 4 for the numbering). This hybrid variant is shown in SEQ ID NO: 2 and is used in the examples below, and is referred to as LE429.
15
20
25

Also, LE174 or LE429 (SEQ ID NO: 2) or *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4 comprising one or more of the following mutations may be used as backbone (using SEQ ID NO: 4 for the numbering of the mutations):

E119C;
S130C;
D124C;
R127C;

A62all possible amino acid residues;
S85all possible amino acid residues;
N96all possible amino acid residues;
V129all possible amino acid residues;
A269all possible amino acid residues;
A378all possible amino acid residues;
S148all possible amino acid residues, in particular S148N;
E211all possible amino acid residues, in particular E211Q;
N188all possible amino acid residues, in particular N188S, N188P
M197all possible amino acid residues, in particular M197T,
M197A, M197C, M197I, M197L, M197Y, M197F, M197I;
W138all possible amino acid residues, in particular W138Y;
D207all possible amino acid residues, in particular D207Y;
H133all possible amino acid residues, in particular H133Y;
H205all possible amino acid residues, in particular H205H,
H205C, H205R;
S187all possible amino acid residues, in particular S187D;
A210all possible amino acid residues, in particular A210S,
A210T;
H405all possible amino acid residues, in particular H405D;
K176all possible amino acid residues, in particular K176R;
F279all possible amino acid residues, in particular F279Y;
Q298all possible amino acid residues, in particular Q298H;
G299all possible amino acid residues, in particular G299R;
L308all possible amino acid residues, in particular L308F;
T412all possible amino acid residues, in particular T412A;

Further, *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4 comprising at least one of the following mutations may be used as backbone:

M15all possible amino acid residues;
A33all possible amino acid residues;

When using LE429 (shown in SEQ ID NO: 2) as the backbone (i.e., as the parent Termamyl-like alpha-amylase) by combining LE174 with the mutation I201F (SEQ ID NO: 4 numbering), the mutations/alterations, in particular substitutions, deletions and insertions, may according to the invention be made in one or more of the following positions to improve the reduced ability

to cleave a substrate close to the branching point, and to improve substrate specificity and/or improved specific activity: W13, G48, T49, S50, Q51, A52, D53, V54, G57, G107, G108, A111, S168, M197 (using the SEQ ID NO: 4 numbering)

5 wherein (a) the alteration(s) are independently

(i) an insertion of an amino acid downstream of the amino acid which occupies the position,

(ii) a deletion of the amino acid which occupies the position, or

10 (iii) a substitution of the amino acid which occupies the position with a different amino acid,

(b) the variant has alpha-amylase activity and (c) each position corresponds to a position of the amino acid sequence of the parent Thermamyl-like alpha-amylase having the amino acid sequence of SEQ ID NO: 4.

In a preferred embodiment a variant of the invention comprises at least one mutation in a position corresponding to the following mutations in the amino acid sequence shown in SEQ ID NO: 4:

20 V54N, A52S, A52S+V54N, T49L, T49+G107A, A52S+V54N+T49L+G107A, A52S+V54N+T49L, G107A, Q51R, Q51R+A52S, A52N; or

T49F+G107A, T49V+G107A, T49D+G107A, T49Y+G107A, T49S+G107A, T49N+G107A, T49I+G107A, T49L+A52S+G107A, T49L+A52T+G107A, T49L+A52F+G107A, T49L+A52L+G107A, T49L+A52I+G107A,

25 T49L+A52V+G107A; or

T49V, T49I, T49D, T49N, T49S, T49Y, T49F, T49W, T49M, T49E, T49Q, T49K, T49R, A52T, A52L, A52I, A52V, A52M, A52F, A52Y, A52W, V54M, G107V, G07I, G107L, G107C.

In a preferred embodiment a variant of the invention comprises at least one mutation in a position corresponding to the following mutations in the amino acid sequence shown in SEQ ID NO: 4:

30 W13F, L, I, V, Y, A;

G48A, V, S, T, I, L;

35 *48aD or *48aY (i.e., insertion of D or Y);

T49X;

*49aX (i.e., insertion of any amino acid residue)

G50X, in particular D,Y,L,T,V,I;
Q51R,K;
A52X, in particular A52S,N,T,P,L,I,V;
D53E,Q,Y,I,N,S,T,V,L;
5 V54X, in particular V54I,N,W,Y,F,L;
G57S,A,V,L,I,F,Y,T;
G107X, in particular G107A,V,S,T,I,L,C;
G108X, in particular G108A,V,S,T,I,L;
A111V,I,L;
10 S168Y;
M197X, in particular Y,F,L,I,T,A,G.

In a preferred embodiment a variant of the invention comprises at least one mutation in a position corresponding to the following mutations in the amino acid sequence shown in SEQ 15 ID NO: 4:

T49X+A52X+V54N/I/L/Y/F/W+G107A, and may further comprise G108A.

In a preferred embodiment a variant of the invention comprises at least one mutation in a position corresponding to the following mutations in the amino acid sequence shown in SEQ 20 ID NO: 4:

T49L+G107A;
T49I+G107A;
T49L+G107A+V54I;
T49I+G107A+V54I;
25 A52S+V54N+T49L+G107A;
A52S+V54I+T49L+G107A;
A52S+T49L+G107A;
A52T+T49L+G107A;
A52S+V54N+T49I+G107A;
30 A52S+V54I+T49I+G107A;
A52S+T49I+G107A;
T49L+G108A;
T49I+G108A;
T49L+G108A+V54I;
35 T49I+G108A+V54I.

General mutations in variants of the invention

It may be preferred that a variant of the invention comprises one or more modifications in addition to those outlined above. Thus, it may be advantageous that one or more proline residues present in the part of the alpha-amylase variant which is modified is/are replaced with a non-proline residue which may be any of the possible, naturally occurring non-proline residues, and which preferably is an alanine, glycine, serine, threonine, valine or leucine.

Analogously, it may be preferred that one or more cysteine residues present among the amino acid residues with which the parent alpha-amylase is modified is/are replaced with a non-cysteine residue such as serine, alanine, threonine, glycine, valine or leucine.

Furthermore, a variant of the invention may - either as the only modification or in combination with any of the above outlined modifications - be modified so that one or more Asp and/or Glu present in an amino acid fragment corresponding to the amino acid fragment 185-209 of SEQ ID NO. 4 is replaced by an Asn and/or Gln, respectively. Also of interest is the replacement, in the Termamyl-like alpha-amylase, of one or more of the Lys residues present in an amino acid fragment corresponding to the amino acid fragment 185-209 of SEQ ID NO: 4 by an Arg.

It will be understood that the present invention encompasses variants incorporating two or more of the above outlined modifications.

Furthermore, it may be advantageous to introduce point-mutations in any of the variants described herein.

30 Methods for preparing alpha-amylase variants

Several methods for introducing mutations into genes are known in the art. After a brief discussion of the cloning of alpha-amylase-encoding DNA sequences, methods for generating mutations at specific sites within the alpha-amylase-encoding sequence will be discussed.

Cloning a DNA sequence encoding an alpha-amylase

The DNA sequence encoding a parent alpha-amylase may be isolated from any cell or microorganism producing the alpha-amylase in question, using various methods well known in the art. First, a genomic DNA and/or cDNA library should be constructed using chromosomal DNA or messenger RNA from the organism that produces the alpha-amylase to be studied. Then, if the amino acid sequence of the alpha-amylase is known, homologous, labelled oligonucleotide probes may be synthesized and used to identify alpha-amylase-encoding clones from a genomic library prepared from the organism in question. Alternatively, a labelled oligonucleotide probe containing sequences homologous to a known alpha-amylase gene could be used as a probe to identify alpha-amylase-encoding clones, using hybridization and washing conditions of lower stringency.

Yet another method for identifying alpha-amylase-encoding clones would involve inserting fragments of genomic DNA into an expression vector, such as a plasmid, transforming alpha-amylase-negative bacteria with the resulting genomic DNA library, and then plating the transformed bacteria onto agar containing a substrate for alpha-amylase, thereby allowing clones expressing the alpha-amylase to be identified.

Alternatively, the DNA sequence encoding the enzyme may be prepared synthetically by established standard methods, e.g., the phosphoroamidite method described by S.L. Beaucage and M.H. Caruthers (1981) or the method described by Matthes et al. (1984). In the phosphoroamidite method, oligonucleotides are synthesized, e.g., in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors.

Finally, the DNA sequence may be of mixed genomic and synthetic origin, mixed synthetic and cDNA origin or mixed genomic and cDNA origin, prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate, the fragments corresponding to various parts of the entire DNA sequence), in accordance with standard techniques. The DNA sequence may also be prepared by polymerase chain reaction (PCR) using specific primers, for instance as described in US 4,683,202 or R.K. Saiki

et al. (1988).

Site-directed mutagenesis

Once an alpha-amylase-encoding DNA sequence has been isolated, and desirable sites for mutation identified, mutations may be introduced using synthetic oligonucleotides. These oligonucleotides contain nucleotide sequences flanking the desired mutation sites; mutant nucleotides are inserted during oligonucleotide synthesis. In a specific method, a single-stranded gap of DNA, bridging the alpha-amylase-encoding sequence, is created in a vector carrying the alpha-amylase gene. Then the synthetic nucleotide, bearing the desired mutation, is annealed to a homologous portion of the single-stranded DNA. The remaining gap is then filled in with DNA polymerase I (Klenow fragment) and the construct is ligated using T4 ligase. A specific example of this method is described in Morinaga et al. (1984). US 4,760,025 disclose the introduction of oligonucleotides encoding multiple mutations by performing minor alterations of the cassette. However, an even greater variety of mutations can be introduced at any one time by the Morinaga method, because a multitude of oligonucleotides, of various lengths, can be introduced.

Another method for introducing mutations into alpha-amylase-encoding DNA sequences is described in Nelson and Long (1989). It involves the 3-step generation of a PCR fragment containing the desired mutation introduced by using a chemically synthesized DNA strand as one of the primers in the PCR reactions. From the PCR-generated fragment, a DNA fragment carrying the mutation may be isolated by cleavage with restriction endonucleases and reinserted into an expression plasmid.

Random Mutagenesis

Random mutagenesis is suitably performed either as localised or region-specific random mutagenesis in at least three parts of the gene translating to the amino acid sequence shown in question, or within the whole gene.

The random mutagenesis of a DNA sequence encoding a parent

alpha-amylase may be conveniently performed by use of any method known in the art.

In relation to the above, a further aspect of the present invention relates to a method for generating a variant of a parent alpha-amylase, e.g., wherein the variant exhibits a reduced capability of cleaving an oligo-saccharide substrate close to the branching point, and further exhibits improved substrate specificity and/or improved specific activity relative to the parent, the method:

- 10 (a) subjecting a DNA sequence encoding the parent alpha-amylase to random mutagenesis,
- (b) expressing the mutated DNA sequence obtained in step (a) in a host cell, and
- 15 (c) screening for host cells expressing an alpha-amylase variant which has an altered property (i.e., thermal stability) relative to the parent alpha-amylase.

Step (a) of the above method of the invention is preferably performed using doped primers. For instance, the random mutagenesis may be performed by use of a suitable physical or 20 chemical mutagenizing agent, by use of a suitable oligonucleotide, or by subjecting the DNA sequence to PCR generated mutagenesis. Furthermore, the random mutagenesis may be performed by use of any combination of these mutagenizing agents. The mutagenizing agent may, e.g., be one, which induces transitions, transversions, inversions, scrambling, deletions, 25 and/or insertions.

Examples of a physical or chemical mutagenizing agent suitable for the present purpose include ultraviolet (UV) ir-radiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate (EMS), sodium bisulphite, formic acid, and nucleotide analogues. When such agents are used, the mutagenesis is typically 30 performed by incubating the DNA sequence encoding the parent enzyme to be mutagenized in the presence of the mutagenizing agent of choice under suitable conditions for the mutagenesis to take place, and selecting for mutated DNA having the desired properties. When the mutagenesis is performed by the use of an

oligonucleotide, the oligonucleotide may be doped or spiked with the three non-parent nucleotides during the synthesis of the oligonucleotide at the positions, which are to be changed. The doping or spiking may be done so that codons for unwanted amino acids are avoided. The doped or spiked oligonucleotide can be incorporated into the DNA encoding the alpha-amylase enzyme by any published technique, using e.g., PCR, LCR or any DNA polymerase and ligase as deemed appropriate. Preferably, the doping is carried out using "constant random doping", in which the percentage of wild type and mutation in each position is predefined. Furthermore, the doping may be directed toward a preference for the introduction of certain nucleotides, and thereby a preference for the introduction of one or more specific amino acid residues. The doping may be made, e.g., so as to allow for the introduction of 90% wild type and 10% mutations in each position. An additional consideration in the choice of a doping scheme is based on genetic as well as protein-structural constraints. The doping scheme may be made by using the DOPE program, which, inter alia, ensures that introduction of stop codons is avoided. When PCR-generated mutagenesis is used, either a chemically treated or non-treated gene encoding a parent alpha-amylase is subjected to PCR under conditions that increase the mis-incorporation of nucleotides (Deschler 1992; Leung et al., Technique, Vol.1, 1989, pp. 11-15). A mutator strain of *E. coli* (Fowler et al., Molec. Gen. Genet., 133, 1974, pp. 179-191), *S. cerevisiae* or any other microbial organism may be used for the random mutagenesis of the DNA encoding the alpha-amylase by, e.g., transforming a plasmid containing the parent glycosylase into the mutator strain, growing the mutator strain with the plasmid and isolating the mutated plasmid from the mutator strain. The mutated plasmid may be subsequently transformed into the expression organism. The DNA sequence to be mutagenized may be conveniently present in a genomic or cDNA library prepared from an organism expressing the parent alpha-amylase. Alternatively, the DNA sequence may be present on a suitable vector such as a plasmid or a bacteriophage, which as such may be incubated with or

otherwise exposed to the mutagenising agent. The DNA to be mutagenized may also be present in a host cell either by being integrated in the genome of said cell or by being present on a vector harboured in the cell. Finally, the DNA to be mutagenized may be in isolated form. It will be understood that the DNA sequence to be subjected to random mutagenesis is preferably a cDNA or a genomic DNA sequence. In some cases it may be convenient to amplify the mutated DNA sequence prior to performing the expression step b) or the screening step c). Such amplification may be performed in accordance with methods known in the art, the presently preferred method being PCR-generated amplification using oligonucleotide primers prepared on the basis of the DNA or amino acid sequence of the parent enzyme. Subsequent to the incubation with or exposure to the mutagenising agent, the mutated DNA is expressed by culturing a suitable host cell carrying the DNA sequence under conditions allowing expression to take place. The host cell used for this purpose may be one which has been transformed with the mutated DNA sequence, optionally present on a vector, or one which was carried the DNA sequence encoding the parent enzyme during the mutagenesis treatment. Examples of suitable host cells are the following: gram positive bacteria such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentinus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus*, *Bacillus megaterium*, *Bacillus thuringiensis*, *Streptomyces lividans* or *Streptomyces murinus*; and gram-negative bacteria such as *E. coli*. The mutated DNA sequence may further comprise a DNA sequence encoding functions permitting expression of the mutated DNA sequence.

Localised random mutagenesis

The random mutagenesis may be advantageously localised to a part of the parent alpha-amylase in question. This may, e.g., be advantageous when certain regions of the enzyme have been identified to be of particular importance for a given property of the enzyme, and when modified are expected to result in a variant having improved properties. Such regions may normally be identified when the tertiary structure of the parent enzyme has been elucidated and related to the function of the enzyme.

The localised, or region-specific, random mutagenesis is conveniently performed by use of PCR generated mutagenesis techniques as described above or any other suitable technique known in the art. Alternatively, the DNA sequence encoding the part of the DNA sequence to be modified may be isolated, e.g., by insertion into a suitable vector, and said part may be subsequently subjected to mutagenesis by use of any of the mutagenesis methods discussed above.

Alternative methods of providing alpha-amylase variants

Alternative methods for providing variants of the invention include gene-shuffling method known in the art including the methods e.g., described in WO 95/22625 (from Affymax Technologies N.V.) and WO 96/00343 (from Novo Nordisk A/S).

Expression of alpha-amylase variants

According to the invention, a DNA sequence encoding the variant produced by methods described above, or by any alternative methods known in the art, can be expressed, in enzyme form, using an expression vector which typically includes control sequences encoding a promoter, operator, ribosome binding site, translation initiation signal, and, optionally, a repressor gene or various activator genes.

The recombinant expression vector carrying the DNA sequence encoding an alpha-amylase variant of the invention may be any vector, which may conveniently be subjected to recombinant DNA procedures, and the choice of vector will often depend on the

host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e., a vector, which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, a bacteriophage or an extrachromosomal element, minichromosome or an artificial chromosome. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

In the vector, the DNA sequence should be operably connected to a suitable promoter sequence. The promoter may be any DNA sequence, which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell. Examples of suitable promoters for directing the transcription of the DNA sequence encoding an alpha-amylase variant of the invention, especially in a bacterial host, are the promoter of the lac operon of *E.coli*, the *Streptomyces coelicolor* agarase gene *dagA* promoters, the promoters of the *Bacillus licheniformis* alpha-amylase gene (*amyL*), the promoters of the *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), the promoters of the *Bacillus amyloliquefaciens* alpha-amylase (*amyQ*), the promoters of the *Bacillus subtilis* *xylA* and *xylB* genes etc. For transcription in a fungal host, examples of useful promoters are those derived from the gene encoding *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral alpha-amylase, *A. niger* acid stable alpha-amylase, *A. niger* glucoamylase, *Rhizomucor miehei* lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase or *A. nidulans* acetamidase.

The expression vector of the invention may also comprise a suitable transcription terminator and, in eukaryotes, polyadenylation sequences operably connected to the DNA sequence encoding the alpha-amylase variant of the invention. Termination and polyadenylation sequences may suitably be derived from the same sources as the promoter.

The vector may further comprise a DNA sequence enabling the

vector to replicate in the host cell in question. Examples of such sequences are the origins of replication of plasmids pUC19, pACYC177, pUB110, pE194, pAMB1 and pIJ702.

The vector may also comprise a selectable marker, e.g., a gene the product of which complements a defect in the host cell, such as the *dal* genes from *B. subtilis* or *B. licheniformis*, or one which confers antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracyclin resistance. Furthermore, the vector may comprise *Aspergillus* selection markers such as *amds*, *argB*, *niaD* and *sC*, a marker giving rise to hygromycin resistance, or the selection may be accomplished by co-transformation, e.g., as described in WO 91/17243.

While intracellular expression may be advantageous in some respects, e.g., when using certain bacteria as host cells, it is generally preferred that the expression is extracellular. In general, the *Sacillus* alpha-amylases mentioned herein comprise a pre-region permitting secretion of the expressed protease into the culture medium. If desirable, this pre-region may be replaced by a different preregion or signal sequence, conveniently accomplished by substitution of the DNA sequences encoding the respective preregions.

The procedures used to ligate the DNA construct of the invention encoding an alpha-amylase variant, the promoter, terminator and other elements, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor, 1989).

The cell of the invention, either comprising a DNA construct or an expression vector of the invention as defined above, is advantageously used as a host cell in the recombinant production of an alpha-amylase variant of the invention. The cell may be transformed with the DNA construct of the invention encoding the variant, conveniently by integrating the DNA construct (in one or more copies) in the host chromosome. This integration is generally considered to be an advantage as the DNA sequence is more likely to be stably maintained in the cell. Integration of

the DNA constructs into the host chromosome may be performed according to conventional methods, e.g., by homologous or heterologous recombination. Alternatively, the cell may be transformed with an expression vector as described above in connection with the different types of host cells.

The cell of the invention may be a cell of a higher organism such as a mammal or an insect, but is preferably a microbial cell, e.g., a bacterial or a fungal (including yeast) cell.

Examples of suitable bacteria are gram-positive bacteria such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus*, *Bacillus megaterium*, *Bacillus thuringiensis*, or *Streptomyces lividans* or *Streptomyces murinus*, or grammegative bacteria such as *E.coli*. The transformation of the bacteria may, for instance, be effected by protoplast transformation or by using competent cells in a manner known *per se*.

The yeast organism may favourably be selected from a species of *Saccharomyces* or *Schizosaccharomyces*, e.g., *Saccharomyces cerevisiae*. The filamentous fungus may advantageously belong to a species of *Aspergillus*, e.g., *Aspergillus oryzae* or *Aspergillus niger*. Fungal cells may be transformed by a process involving protoplast formation and transformation of the protoplasts followed by regeneration of the cell wall in a manner known *per se*. A suitable procedure for transformation of *Aspergillus* host cells is described in EP 238 023.

In yet a further aspect, the present invention relates to a method of producing an alpha-amylase variant of the invention, which method comprises cultivating a host cell as described above under conditions conducive to the production of the variant and recovering the variant from the cells and/or culture medium.

The medium used to cultivate the cells may be any conventional medium suitable for growing the host cell in question and obtaining expression of the alpha-amylase variant of the invention. Suitable media are available from commercial suppliers or

may be prepared according to published recipes (e.g., as described in catalogues of the American Type Culture Collection).

The alpha-amylase variant secreted from the host cells may conveniently be recovered from the culture medium by well-known procedures, including separating the cells from the medium by centrifugation or filtration, and precipitating proteinaceous components of the medium by means of a salt such as ammonium sulphate, followed by the use of chromatographic procedures such as ion exchange chromatography, affinity chromatography, or the like.

Industrial applications

The alpha-amylase variants of this invention possess valuable properties allowing for a variety of industrial applications. In particular, enzyme variants of the invention are applicable as a component in washing, dishwashing and hard surface cleaning detergent compositions. Numerous variants are particularly useful in the production of sweeteners and ethanol, e.g., fuel, drinking or industrial ethanol, from starch, and/or for textile desizing. Conditions for conventional starch-conversion processes, including starch liquefaction and/or saccharification processes, are described in, e.g., US 3,912,580 and in EP patent publications Nos. 252 730 and 63 909.

Production of sweeteners from starch:

A "traditional" process for conversion of starch to fructose syrups normally consists of three consecutive enzymatic processes, viz. a liquefaction process followed by a saccharification process and an isomerization process. During the liquefaction process, starch is degraded to dextrans by an alpha-amylase (e.g., TermamylTM) at pH values between 5.5 and 6.2 and at temperatures of 95-160°C for a period of approx. 2 hours. In order to ensure optimal enzyme stability under these conditions, 1 mM of calcium is added (40 ppm free calcium ions).

After the liquefaction process the dextrans are converted into dextrose by addition of a glucoamylase (e.g., AMGTM) and a

debranching enzyme, such as an isoamylase or a pullulanase (e.g., Promozyme™). Before this step the pH is reduced to a value below 4.5, maintaining the high temperature (above 95°C), and the liquefying alpha-amylase activity is denatured. The temperature is lowered to 60°C, and glucoamylase and debranching enzyme are added. The saccharification process proceeds for 24-72 hours.

After the saccharification process the pH is increased to a value in the range of 6-8, preferably pH 7.5, and the calcium is removed by ion exchange. The dextrose syrup is then converted into high fructose syrup using, e.g., an immobilized glucoseisomerase (such as Sweetzyme™).

At least one enzymatic improvement of this process could be envisaged: Reduction of the calcium dependency of the liquefying alpha-amylase. Addition of free calcium is required to ensure adequately high stability of the alpha-amylase, but free calcium strongly inhibits the activity of the glucoseisomerase and needs to be removed, by means of an expensive unit operation, to an extent, which reduces the level of free calcium to below 3-5 ppm. Cost savings could be obtained if such an operation could be avoided and the liquefaction process could be performed without addition of free calcium ions.

To achieve that, a less calcium-dependent Termamyl-like alpha-amylase which is stable and highly active at low concentrations of free calcium (< 40 ppm) is required. Such a Termamyl-like alpha-amylase should have a pH optimum at a pH in the range of 4.5-6.5, preferably in the range of 4.5-5.5.

The invention also relates to a composition comprising a mixture of one or more variants of the invention derived from (as the parent Termamyl-like alpha-amylase) the *B. stearothermophilus* alpha-amylase having the sequence shown in SEQ ID NO: 8 and a Termamyl-like alpha-amylase derived from the *B. licheniformis* alpha-amylase having the sequence shown in SEQ ID NO: 4.

Further, the invention also relates to a composition comprising a mixture of one or more variants according the

invention derived from (as the parent Termamyl-like alpha-amylase) the *B. stearothermophilus* alpha-amylase having the sequence shown in SEQ ID NO: 8 and a hybrid alpha-amylase comprising a part of the *B. amyloliquefaciens* alpha-amylase shown in SEQ ID NO: 6 and a part of the *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4. The latter mentioned hybrid Termamyl-like alpha-amylase comprises the 445 C-terminal amino acid residues of the *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4 and the 37 N-terminal amino acid residues of the alpha-amylase derived from *B. amyloliquefaciens* shown in SEQ ID NO: 6. Said latter mentioned hybrid alpha-amylase may suitably comprise the following mutations: H156Y+A181T+N196F+A209V+Q264S (using the numbering in SEQ ID NO: 4). Preferably, said latter mentioned hybrid alpha-amylase may suitably comprise the following mutations: H156Y+A181T+N196F+A209V+Q264S+I261F (using the SEQ ID NO: 4 numbering). In the examples below said last-mentioned parent hybrid Termamyl-like alpha-amylase referred to as LE429 (shown in SEQ ID NO: 2) is used for preparing variants of the invention, which variants may be used in compositions of the invention.

An alpha-amylase variant of the invention or a composition of the invention may in an aspect of the invention be used for starch liquefaction, in detergent composition, such as laundry, dish wash compositions and hard surface cleaning, ethanol production, such as fuel, drinking and industrial ethanol production, desizing of textile, fabric and garments.

MATERIALS AND METHODS

Enzymes:

LE174; hybrid alpha-amylase variant:

LE174 is a hybrid Termamyl-like alpha-amylase being identical to the Termamyl sequence, i.e., the *Bacillus licheniformis* alpha-amylase shown in SEQ ID NO: 4, except that the N-terminal 35 amino acid residues (of the mature protein) has been replaced by the N-terminal 33 residues of BAN (mature protein), i.e., the *Bacillus amyloliquefaciens* alpha-amylase shown in SEQ ID NO: 6, which further have following mutations:

H156Y+A181T+N190F+A209V+Q264S (SEQ ID NO: 4).

LE429 hybrid alpha-amylase variant:

LE429 is a hybrid Termamyl-like alpha-amylase being identical to the Termamyl sequence, i.e., the *Bacillus licheniformis* alpha-amylase shown in SEQ ID NO: 4, except that the N-terminal 35 amino acid residues (of the mature protein) has been replaced by the N-terminal 33 residues of BAN (mature protein), i.e., the *Bacillus amyloliquefaciens* alpha-amylase shown in SEQ ID NO: 6, which further have following mutations: H156Y+A181T+N190F+A209V+Q264S+I201F (SEQ ID NO: 4). LE429 is shown as SEQ ID NO: 2 and was constructed by SOE-PCR (Higuchi et al. 1989, Nucleic Acids Research 16:7351).

Dextrozyme™ E: a balanced mixture of glucoamylase (AMG) and pullulanase obtainable from selected strains of *Aspergillus niger* and *Bacillus deramificans* (available from Novo Nordisk A/S)

Fermentation and purification of alpha-amylase variants

A *B. subtilis* strain harbouring the relevant expression plasmid is streaked on an LB-agar plate with 10 micro g/ml kanamycin from -80°C stock, and grown overnight at 37°C. The colonies are transferred to 100 ml SPX media supplemented with 10 micro g/ml kanamycin in a 500 ml shaking flask.

Composition of SPX medium:

Potato starch	100	g/l
Barley flour	50	g/l
BAN 5000 SKB	0.1	g/l
Sodium caseinate	10	g/l
Soy Bean Meal	20	g/l
Na ₂ HPO ₄ , 12 H ₂ O	9	g/l
pluronic™	0.1	g/l

The culture is shaken at 37°C at 270 rpm for 5 days.

Cells and cell debris are removed from the fermentation broth by centrifugation at 4500 rpm in 20-25 minutes. Afterwards the supernatant is filtered to obtain a completely clear

solution. The filtrate is concentrated and washed on an UF-filter (10000 cut off membrane) and the buffer is changed to 20mM Acetate pH 5.5. The UF-filtrate is applied on a S-sepharose F.F. and elution is carried out by step elution with 0.2M NaCl in the same buffer. The eluate is dialysed against 10mM Tris, pH 9.0 and applied on a Q-sepharose F.F. and eluted with a linear gradient from 0-0.3M NaCl over 6 column volumes. The fractions that contain the activity (measured by the Phadebas assay) are pooled, pH was adjusted to pH 7.5 and remaining color was removed by a treatment with 0.5% w/vol. active coal in 5 minutes.

Activity determination - (KNU)

One Kilo alpha-amylase Unit (1 KNU) is the amount of enzyme which breaks down 5.26 g starch (Merck, Amylum Solubile, Erg. B 6, Batch 9947275) per hour in Novo Nordisk's standard method for determination of alpha-amylase based upon the following condition:

Substrate	soluble starch
Calcium content in solvent	0.0043 M
Reaction time	7-20 minutes
Temperature	37°C
pH	5.6

Detailed description of Novo Nordisk's analytical method (AF 28 S) is available on request.

Assay for Alpha-Amylase Activity

Alpha-Amylase activity is determined by a method employing Phadebas® tablets as substrate. Phadebas tablets (Phadebas® Amylase Test, supplied by Pharmacia Diagnostic) contain a cross-linked insoluble blue-coloured starch polymer, which has been mixed with bovine serum albumin and a buffer substance and tabletted.

For every single measurement one tablet is suspended in a tube containing 5 ml 50 mM Britton-Robinson buffer (50 mM acetic acid, 50 mM phosphoric acid, 50 mM boric acid, 0.1 mM CaCl₂, pH adjusted to the value of interest with NaOH). The test is

performed in a water bath at the temperature of interest. The alpha-amylase to be tested is diluted in x ml of 50 mM Britton-Robinson buffer. 1 ml of this alpha-amylase solution is added to the 5 ml 50 mM Britton-Robinson buffer. The starch is hydrolysed by the alpha-amylase giving soluble blue fragments. The absorbance of the resulting blue solution, measured spectrophotometrically at 620 nm, is a function of the alpha-amylase activity.

It is important that the measured 620 nm absorbance after 10 or 15 minutes of incubation (testing time) is in the range of 0.2 to 2.0 absorbance units at 620 nm. In this absorbance range there is linearity between activity and absorbance (Lambert-Beer law). The dilution of the enzyme must therefore be adjusted to fit this criterion. Under a specified set of conditions (temp., pH, reaction time, buffer conditions) 1 mg of a given alpha-amylase will hydrolyse a certain amount of substrate and a blue colour will be produced. The colour intensity is measured at 620 nm. The measured absorbance is directly proportional to the specific activity (activity/mg of pure alpha-amylase protein) of the alpha-amylase in question under the given set of conditions.

Determining Specific Activity

The specific activity is determined using the Phadebas assay (Pharmacia) as activity/mg enzyme.

25

Measuring the pH activity profile (pH stability)

The variant is stored in 20 mM TRIS pH 7.5, 0.1 mM, CaCl₂, and tested at 30°C, 50 mM Britton-Robinson, 0.1 mM CaCl₂. The pH activity is measured at pH 4.0, 4.5, 5.0, 5.5, 6.0, 7.0, 8.0, 9.0, 9.5, 10, and 10.5, using the Phadebas assay described above.

Determination Of AGU Activity and As AGU/mg

One Novo Amyloglucosidase Unit (AGU) is defined as the amount of enzyme, which hydrolyzes 1 micromole maltose per minute at 37°C and pH 4.3. A detailed description of the

analytical method (AEL-SM-0131) is available on request from Novo Nordisk.

The activity is determined as AGU/ml by a method modified after (AEL-SM-0131) using the Glucose GOD-Perid kit from Boehringer Mannheim, 124036. Standard: AMG-standard, batch 7-1195, 195 AGU/ml.

375 microl substrate (1% maltose in 50 mM Sodium acetate, pH 4.3) is incubated 5 minutes at 37°C. 25 microl enzyme diluted in sodium acetate is added. The reaction is stopped after 10 minutes by adding 100 microl 0.25 M NaOH. 20 microl is transferred to a 96 well microtitre plate and 200 microl GOD-Perid solution is added. After 30 minutes at room temperature, the absorbance is measured at 650 nm and the activity calculated in AGU/ml from the AMG-standard.

18 The specific activity in AGU/mg is then calculated from the activity (AGU/ml) divided with the protein concentration (mg/ml).

EXAMPLES

EXAMPLE 1

Construction of Termamyl variants in accordance with the invention

Termamyl (*B. licheniformis* alpha-amylase SEQ ID NO: 4) is expressed in *B. subtilis* from a plasmid denoted pDN1528. This plasmid contains the complete gene encoding Termamyl, *amyL*, the expression of which is directed by its own promoter. Further, the plasmid contains the origin of replication, *ori*, from plasmid pUB110 and the *cat* gene from plasmid pC194 conferring resistance towards chloramphenicol. pDN1528 is shown in Fig. 9 of WO 96/23874.

A specific mutagenesis vector containing a major part of the coding region of SEQ ID NO: 3 was prepared. The important features of this vector, denoted pJSENI, include an origin of replication derived from the pUC plasmids, the *cat* gene conferring resistance towards chloramphenicol, and a frameshift-containing version of the *bla* gene, the wild type of which normally confers resistance towards ampicillin (*amp*^R)

phenotype). This mutated version results in an amp^R phenotype. The plasmid pJeEN1 is shown in Fig. 10 of WO 96/23874, and the E. coli origin of replication, ori, bla, cat, the 5'-truncated version of the Termamyl amylase gene, and selected restriction sites are indicated on the plasmid.

Mutations are introduced in amyL by the method described by Deng and Nickoloff (1992, Anal. Biochem. 200, pp. 81-88) except that plasmids with the "selection primer" (primer #6616; see below) incorporated are selected based on the amp^R phenotype of transformed E. coli cells harboring a plasmid with a repaired bla gene, instead of employing the selection by restriction enzyme digestion outlined by Deng and Nickoloff. Chemicals and enzymes used for the mutagenesis were obtained from the Chameleon[®] mutagenesis kit from Stratagene (catalogue number 200509).

After verification of the DNA sequence in variant plasmids, the truncated gene, containing the desired alteration, is subcloned into pDN1528 as a PstI-EcoRI fragment and transformed into the protease- and amylase-depleted *Bacillus subtilis* strain SHA273 (described in WO92/11357 and WO95/10603) in order to express the variant enzyme.

The Termamyl variant V54W was constructed by the use of the following mutagenesis primer (written 5' to 3', left to right):

PG CTC GTA GCC ACC GCA CCC CCA ATC CGG TTG (SEQ ID NO: 9)

The Termamyl variant AS2W + V54W was constructed by the use of the following mutagenesis primer (written 5' to 3', left to right):

PG CTC GTA GCC ACC GCA CCC CCA ATC CGG GCT CG (SEQ ID NO: 10)

Primer #6616 (written 5' to 3', left to right; P denotes a 5' phosphate):

P CTG TCA CTG GTG AGT ACT CAA CCA AGT C (SEQ ID NO: 11)

The Termamyl variant VS4E was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGG TCG TAG GCA CCG TAG CCC TCA TCC GCT TG (SEQ ID NO: 12)

The Termamyl variant V54M was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGC TCG TAG GCA CCG TAG CCC ATA TCC GCT TG (SEQ ID NO: 13)

The Termamyl variant V54I was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGC TCG TAG GCA CCG TAG CCA ATA TCC GCT TG (SEQ ID NO: 14)

The Termamyl variants Y290E and Y290K were constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGC AGC ATG GAA CTG CTY ATG AAG AGG CAC GTC AAA C (SEQ ID NO: 15)

Y represents an equal mixture of C and T. The presence of a codon encoding either Glutamate or Lysine in position 290 was verified by DNA sequencing.

The Termamyl variant N190F was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PCA TAG TTG CCG AAT TCA TTG GAA ACT TCC C (SEQ ID NO: 16)

The Termamyl variant N188P+N190F was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PCA TAG TTG CCG AAT TCA GGG GAA ACT TCC CAA TC (SEQ ID NO: 17)

The Termamyl variant H140K+H142D was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGC CGC CCC CGG GAA ATC AAA TTT TGT CCA GGC TTT AAT TAG (SEQ ID NO: 18)

The Termamyl variant H156Y was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PCA AAA TGG TAC CAA TAC CAC TTA AAA TCG CTG (SEQ ID NO: 19)

The Termamyl variant A181T was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGT TCC CAA TCC CAA GTC TTC CCT TGA AAC (SEQ ID NO: 20)

The Termamyl variant A209V was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGT AAT TTC TGC TAC GAC GTC AGG ATG GTC ATA ATC (SEQ ID NO: 21)

The Termamyl variant Q264S was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGC CCC AAG TCA TTC GAC CAG TAC TCA GCT ACC GTA AAC (SEQ ID NO: 22)

The Termamyl variant S187D was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGC CGT TTT CAT TGT CGA CTT CCC AAT CCC (SEQ ID NO: 23)

The Termamyl variant DELTA(K370-G371-D372) (i.e., deleted of amino acid residues nos. 370, 371 and 372) was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGG AAT TTC GCG CTG ACT AGT CCC GTA CAT ATC CCC (SEQ ID NO: 24)

The Termamyl variant DELTA(D372-S373-Q374) was constructed by the use of the following mutagenesis primer (written 5'-3', left to right):

PGG CAG GAA TTT CGC GAC CTT TCG TCC CGT ACA TAT C (SEQ ID NO: 25)

The Termamyl variants A181T and A209V were combined to A181T-A209V by digesting the A181T containing pDN1528-like plasmid (i.e., pDN1528 containing within amyL the mutation resulting in the A181T alteration) and the A209V-containing pDN1528-like plasmid (i.e., pDN1528 containing within amyL the mutation resulting in the A209V alteration) with restriction enzyme *Cla*I which cuts the pDN1528-like plasmids twice resulting in a fragment of 1116 bp and the vector-part (i.e. contains the plasmid origin of replication) of 3850 bp. The fragment containing the A209V mutation and the vector part containing the A181T mutation were purified by QIAquick gel extraction kit (purchased from QIAGEN) after separation on an

agarose gel. The fragment and the vector were ligated and transformed into the protease and amylase depleted *Bacillus subtilis* strain referred to above. Plasmid from amy+ (clearing zones on starch containing agar-plates) and chloramphenicol resistant transformants were analysed for the presence of both mutations on the plasmid.

In a similar way as described above, H156Y and A209V were combined utilizing restriction endonucleases Acc65I and EcoRI, giving H156Y+A209V.

H156Y +A209V and A181T+A209V were combined into H156Y+ A181T+A209V by the use of restriction endonucleases Acc65I and HindIII.

The 35 N-terminal residues of the mature part of Termamyl variant H156Y+ A181T+A209V were substituted by the 33 N-terminal residues of the *B. amyloliquefaciens* alpha-amylase (SEQ ID NO: 4) (which in the present context is termed BAN) by a SOE-PCR approach (Higuchi et al. 1988, Nucleic Acids Research 16:7351) as follows:

Primer 19364 (sequence 5'-3'): CCT CAT TCT GCA GCA GCA GCC GTA
AAT GGC ACG CTG (SEQ ID NO: 26)

Primer 19362: CCA GAC GGC AGT AAT ACC GAT ATC CGA TAA ATG TTC
CG (SEQ ID NO: 27)

Primer 19363: CGG ATA TCG GTA TTA CTG CGG TCT CGA TTC (SEQ ID
NO: 28)

Primer 1C: CTC GTC CCA ATC CGT TCC GTC (SEQ ID NO: 29)

A standard PCR, polymerase chain reaction, was carried out using the Pwo thermostable polymerase from Boehringer Mannheim according to the manufacturer's instructions and the temperature cyclus: 5 minutes at 94°C, 25 cycles of (94°C for 30 seconds, 50°C for 45 seconds, 72°C for 1 minute), 72°C for 10 minutes.

An approximately 130 bp fragment was amplified in a first PCR denoted PCR1 with primers 19364 and 19362 on a DNA fragment containing the gene encoding the *B. amyloliquefaciens* alpha-amylase.

An approximately 400 bp fragment was amplified in another PCR denoted PCR2 with primers 19363 and 1C on template

pDN1538.

PCR1 and PCR2 were purified from an agarose gel and used as templates in PCR3 with primers 19364 and 1C, which resulted in a fragment of approximately 520 bp. This fragment thus contains one part of DNA encoding the N-terminus from BAN fused to a part of DNA encoding Termamyl from the 35th amino acid.

The 520 bp fragment was subcloned into a pDN1528-like plasmid (containing the gene encoding Termamyl variant H156Y+ A181T+A209V) by digestion with restriction endonucleases PstI and SacII, ligation and transformation of the *B. subtilis* strain as previously described. The DNA sequence between restriction sites PstI and SacII was verified by DNA sequencing in extracted plasmids from amy+ and chloramphenicol resistant transformants.

The final construct containing the correct N-terminus from BAN and H156Y+ A181T+A209V was denoted BAN(1-35)+ H156Y+ A181T+A209V.

N190F was combined with BAN(1-35)+ H156Y+ A181T+A209V giving BAN(1-35)+ H156Y+ A181T+N190F+A209V by carrying out mutagenesis as described above except that the sequence of amyL in pJcEN1 was substituted by the DNA sequence encoding Termamyl variant BAN(1-35)+ H156Y+ A181T+A209V

Q264S was combined with BAN(1-35)+ H156Y+ A181T+A209V giving BAN(1-35)+ H156Y+ A181T+A209V+Q264S by carrying out mutagenesis as described above except that the sequence of amyL in pJcEN was substituted by the DNA sequence encoding Termamyl variant BAN(1-35)+ H156Y+ A181T+A209V

BAN(1-35)+ H156Y+ A181T+A209V+Q264S and BAN(1-35)+ H156Y+ A181T+N190F+A209V were combined into BAN(1-35)+ H156Y+ A181T+N190F+A209V+Q264S utilizing restriction endonucleases *Bsa*HI (*Bsa*HI site was introduced close to the A209V mutation) and PstI.

I201F was combined with BAN(1-35)+ H156Y+ A181T+N190F+A209V+Q264S giving BAN(1-35)+ H156Y+ A181T+N190F+A209V+Q264S+I201F (SEQ ID NO: 2) by carrying out mutagenesis as described above. The mutagenesis primer AM100

was used, introduced the I201F substitution and removed simultaneously a Cla I restriction site, which facilitates easy pin-pointing of mutants.

5 primer AM100:

5'GATGTATGCCGACTTCGATTAGACC 3' (SEQ ID NO: 30)

EXAMPLE 2

10 Construction of Termamyl-like alpha-amylase variants with an altered cleavage pattern according to the invention

The variant of the thermostable *B. licheniformis* alpha-amylase consisting comprising the 445 C-terminal amino acid residues of the *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4 and the 37 N-terminal amino acid residues of the alpha-amylase derived from *B. amyloliquefaciens* shown in SEQ ID NO: 6, and further comprising the following mutations:

15 H156Y+A181T+N190F+A209V+Q264S+I201F (the construction of this variant is described in Example 1, and the amino acid sequence shown in SEQ ID NO: 2) has a reduced capability of cleaving an substrate close to the branching point.

In an attempt to further improve the reduced capability 20 of cleaving an substrate close to the branching point of said alpha-amylase variant site directed mutagenesis was carried out using the Mega-primer method as described by Sarker and Sommer, 1990 (BioTechniques 8: 404-407):

Construction of LB313: BAN/Termamyl hybrid *
H156Y+A181T+N190F+ A209V+Q264S+V54N:

25 Gene specific primer 27274 and mutagenic primer AM116 are used to amplify by PCR an approximately 440 bp DNA fragment from a pDN1526-like plasmid (harbouring the BAN(I-2S)+H156Y+A181T+N190F+I201F+A209V+Q264S mutations in the gene encoding the amylase from SEQ ID NO: 4).

The 440 bp fragment is purified from an agarose gel and used 30 as a Mega-primer together with primer 113711 in a second PCR carried out on the same template.

The resulting approximately 630 bp fragment is digested

with restriction enzymes EcoR V and Acc65 I and the resulting approximately 370 bp DNA fragment is purified and ligated with the pDN1528-like plasmid digested with the same enzymes. Competent *Bacillus subtilis* SPA273 (amylase and protease low) cells are transformed with the ligation and Chloramphenicol resistant transformants are checked by DNA sequencing to verify the presence of the correct mutations on the plasmid.

Primer 27274:

16 5' CATACTTCCCGAATTCAATTGGAAACTTCCC 3' (SEQ ID NO: 31)

Primer 1B:

5' CCCATTGCTGACCGCTGTTATTCGC 3' (SEQ ID NO: 32)

18 primer AM115:

5' GCCAACGGGATAACGGCTACGGTGC 3' (SEQ ID NO: 33)

Construction of LE314: BAN/Termamyl hybrid + H156Y+A181T+N190F+ A209V+Q264S + AS2S is carried out in a similar way, except that mutagenic primer AM116 is used.

AM116:

5' CAACGAGCCAATCGGACGTGGCTACCG 3' (SEQ ID NO: 34)

26 Construction of LE315: BAN/Termamyl hybrid + H156Y+A181T+N190F+ A209V+Q264S + AS2S+V54N is carried out in a similar way, except that mutagenic primer AM117 is used.

AM117:

30 5' CGAACGGAGCCAATCGGATAACGGCTACGGTGC 3' (SEQ ID NO: 35)

Construction of LE316: BAN/Termamyl hybrid + H156Y+A181T+N190F+ A209V+Q264S + T49L is carried out in a similar way, except that mutagenic primer AM118 is used.

AM118:

35 5' GCATATAACGGACTGAGCCAAGCGG 3' (SEQ ID NO: 36)

Construction LE317: BAN/Termamyl hybrid + H156Y+A181T+N190F+
A209V+Q264S + T49L+G107A is carried our in a similar way,
except that mutagenic primer AM118 and mutagenic primer AM119
are used simultaneously.

AM119:

5' CAACCACAAAGCCCCCGCTGATGCG 3' (SEQ ID NO: 37)

Construction of LE316: BAN/Termamyl hybrid +
H156Y+A181T+N190F+ A209V+Q264S + A52S+V54N+T49L+G107A is
carried our in a similar way, except that mutagenic primer
AM120 and mutagenic primer AM119 are used simultaneously.

AM120:

5' GCATATAACGGACTGAGCCAATCCGATAACGGCTACGGTGC 3' (SEQ ID NO:
38)

Construction of LE 319: BAN/Termamyl hybrid +
H156Y+A181T+N190F+ A209V+Q264S + A52S+V54N+T49L is carried our
in a similar way, except that mutagenic primer AM120 is used.

Construction of LE320: BAN/Termamyl hybrid +
H156Y+A181T+N190F+ A209V+Q264S + G107A is carried our in a
similar way, except that mutagenic primer AM119 is used.

Construction of LE322: BAN/Termamyl hybrid +
H156Y+A181T+N190F+A209V+Q264S + Q61R+A52S is carried our in a
similar way, except that mutagenic primer AM121 is used.

AM121:

5' GAACGAGCCGATCGGACGTGGCTACGG 3' (SEQ ID NO:39)

Construction of LE323: BAN/Termamyl hybrid +
H156Y+A181T+N190F+ A209V+Q264S + A52N is carried our in a
similar way, except that mutagenic primer AM122 is used.

AM122:

5' GAACGAGCCAAACGACGTGGCTACGG 3' (SEQ ID NO: 40)

EXAMPLE 3

Testing of LE429 variants (saccharification)

The standard reaction conditions were:

Substrate concentration	30 % w/w
Temperature	60°C
Initial pH (at 60°C)	5.5
Enzyme dosage	
Glucosamylase	0.18 AGU/g DS
pullulanase	0.06 PUN/g DS
Alpha-amylase	10 micro g enzyme/g DS

Dextroxyme™ S was used to provide glucosamylase and pullulanase activities.

Substrates for saccharification were prepared by dissolving common corn starch in deionized water and adjusting the dry substance to approximately 30% w/w. The pH was adjusted to 5.5 (measured at 60°C), and aliquots of substrate corresponding to 10 g dry weight were transferred to blue cap glass flasks.

The flasks were then placed in a shaking water bath equilibrated at 60°C, and the enzymes added. The pH was readjusted to 5.5 where necessary. The samples were taken after 48 hours of saccharification; the pH was adjusted to about 3.0, and then heated in a boiling water bath for 15 minutes to inactivate the enzymes. After cooling, the samples were treated with approximately 0.1 g mixed bed ion exchange resin (BIO-RAD 501 X8 (D)) for 30 minutes on a rotary mixer to remove salts and soluble N. After filtration, the carbohydrate composition was determined by HPLC. The following results were obtained:

The parent alpha-amylase for the variants is LE429.

Added Alpha-amylase Variants	DP ₁	DP ₂	DP ₃	SPEC. ACT. (NU/mg)
V54N	96.1	1.75	1.18	8200
A52S	95.9	1.80	1.11	18800
A52S+V54N	96.3	1.84	1.02	10000

T49L	96.3	1.77	1.11	12300
T49L+G107A	96.4	1.87	0.72	13600
A52S+V54N+T49L+G107A	80.5	2.55	0.43	10000
A52S+V54N+T49L	95.8	1.76	0.84	8400
G107A	94.4	1.89	1.04	19500
Q51R+A52S	95.9	1.77	1.27	16500
A52N	95.5	1.89	1.56	17600
LEI74 (CONTROL)		95.97	1.877	1.177
		95.8	1.83	1.35

Compared with the control, the presence of an active alpha-amylase variant of the invention during liquefaction results in decreased panose levels (DP₃).

Especially the T49L+G107A variant of LE429 and the A52S+V54N+T49L variant of LE429, respectively, result in a drastically decreased panose level (DP₃). If these alpha-amylase variants are used for starch liquefaction, it will not be necessary to inactivate the enzyme before the commencement of saccharification.

Example 4

Liquefaction and saccharification of LE429 variants

The experiment in Example 3 was repeated for a number of other LE429 variants under the same conditions.

The result is shown below:

Variant/sugar profile	DP1	DP2	DP3	DP4+
T49V+G107A	95.9%	1.72%	1.27%	1.11%
T49Y+G107A	95.3%	1.73%	1.29%	1.65%
T49N+G107A	95.7%	1.64%	1.51%	1.18%
T49L+A52S+G107A	95.7%	1.73%	0.95%	1.67%
T49L+A52T+G107A	95.8%	1.66%	1.03%	1.46%
T49L+A52F+G107A	95.7%	1.69%	1.16%	1.42%
T49L+A52L+G107A	95.5%	1.70%	1.40%	1.38%
T49L+A52I+G107A	95.9%	1.72%	1.31%	1.07%
T49L+A52V+G107A	94.7%	1.69%	1.16%	2.44%

T49L+A52V+G107A+A111V	94.5%	1.75%	0.72%	2.99%
LE429	94.9%	1.71%	1.65%	1.51%

Example 5

The experiment in Example 3 was repeated for a number of LE429 variants, except that the liquefaction was carried out at 95°C, pH 6.0 and the saccharification at 60°C, pH 4.5, 40 ppm CaCl₂, followed by inactivation. The variant referred to below are LE429 variant. The results found are as follows:

10

Variant/sugar profile	DP4+	DP3	DP2	DP1
T49F	1.15	0.92	1.83	96.12
T49D+G107A	0.84	1.03	1.82	96.3
T49I+G107A	0.97	0.64	1.84	96.55
T49L+G107A	0.96	0.81	1.82	96.42
T49L+A52S+G107A	1.37	0.75	1.88	96.01
T49L+A52T+G107A	0.87	0.81	1.8	96.52
T49L+A52F+G107A	0.98	0.83	1.87	96.31
T49V+G107A	0.65	0.8	2.13	96.43
T49Y+G107A	0.83	0.94	1.89	96.35
LE429	1.16	1.21	1.77	96.87

REFERENCES CITED

Klein, C., et al., *Biochemistry* 1992, 31, 8740-8746,

Mizuno, H., et al., *J. Mol. Biol.* (1993) 234, 1262-1283,

8 Chang, C., et al., *J. Mol. Biol.* (1993) 229, 235-238,

Larson, S.B., *J. Mol. Biol.* (1994) 235, 1560-1564,

Lawson, C.L., *J. Mol. Biol.* (1994) 236, 590-600,

Qian, M., et al., *J. Mol. Biol.* (1993) 231, 785-799,

Brady, R.L., et al., *Acta Crystallogr. sect. B*, 47, 527-535,

10 Swift, H.J., et al., *Acta Crystallogr. sect. B*, 47, 535-544

A. Kadziola, Ph.D. Thesis: "An alpha-amylase from Barley and its Complex with a Substrate Analogue Inhibitor Studied by X-ray Crystallography", Department of Chemistry University of Copenhagen 1993

15 MacGregor, E.A., *Food Hydrocolloids*, 1987, Vol.1, No. 5-6, p.

B. Diderichsen and L. Christiansen, Cloning of a maltogenic amylase from *Bacillus stearothermophilus*, *FEMS Microbiol. letters*: 56: pp. 53-60 (1988)

Hudson et al., *Practical Immunology*, Third edition (1989),
20 Blackwell Scientific Publications,

Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor, 1989

S.L. Beaucage and M.H. Caruthers, *Tetrahedron Letters* 32, 1981,
pp. 1859-1869

25 Matthes et al., *The EMBO J.* 3, 1984, pp. 801-805.

R.K. Saiki et al., *Science* 239, 1988, pp. 487-491.

Morinaga et al., (1984, *Biotechnology* 2:646-653)

Nelson and Long, *Analytical Biochemistry* 180, 1989, pp. 147-151

Hunkapiller et al., 1984, *Nature* 310:108-111

30 R. Miguchi, B. Krummel, and R.K. Saiki (1988). A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucl. Acids Res.* 16:7351-7367.

Dubnau et al., 1971, *J. Mol. Biol.* 56, pp. 209-221.

35 Gryczan et al., 1978, *J. Bacteriol.* 134, pp. 318-329.

S.D. Erlich, 1977, *Proc. Natl. Acad. Sci.* 74, pp. 1680-1682.

Boel et al., 1990, *Biochemistry* 29, pp. 6244-6249.

Sarkar and Sommer, 1990, BioTechniques 8, pp. 404-407.

CLAIMS

1. A variant of a parent Ternamyl-like alpha-amylase, comprising an alteration at one or more positions selected from the group of:
W13, G48, T49, S50, Q51, A52, D53, V54, G57, G107, G108, A111, S168, M197, wherein (a) the alteration(s) are independently
 - (i) an insertion of an amino acid downstream of the amino acid which occupies the position,
 - (ii) a deletion of the amino acid which occupies the position, or
 - (iii) a substitution of the amino acid which occupies the position with a different amino acid,

(b) the variant has alpha-amylase activity and (c) each position corresponds to a position of the amino acid sequence of the parent Ternamyl-like alpha-amylase having the amino acid sequence of SEQ ID NO: 4.
2. The variant of claims 1, comprises a mutation in a position corresponding to at least one of the following mutations in the amino acid sequence shown in SEQ ID NO: 4:
V54N, A52S, A52S+V54N, T49L, T49+G107A, A52S+V54N+T49L+G107A, A52S+V54N+T49L, G107A, Q51R, Q51R+A52S, A52N; or
T49F+G107A, T49V+G107A, T49D+G107A, T49Y+G107A, T49S+G107A,
T49N+G107A, T49I+G107A, T49L+A52S+G107A, T49L+A52T+G107A,
T49L+A52F+G107A, T49L+A52L+G107A, T49L+A52I+G107A,
T49L+A52V+G107A; or
T49V, T49I, T49D, T49N, T49S, T49Y, T49F, T49W, T49M, T49E, T49Q, T49K, T49R, A52T, A52L, A52I, A52V, A52M, A52F, A52Y, A52W, V54M, Q107V, G07I, G107L, G107C.
3. The variant of claims 1 or 2, comprising a mutation in a position corresponding to at least one of the following mutations in the amino acid sequence shown in SEQ ID NO: 4:
W13F, L,I,V,Y,A;
G48A,V,S,T,I,L;
*48aD or *48aY (i.e., insertion of U or Y);

T49X;
*49aX (i.e., insertion of any amino acid residue)
S50X, in particular D,Y,L,T,V,I;
Q51R,K;
A52X, in particular A52S,N,T,F,L,I,V;
D53E,Q,Y,I,N,S,T,V,L;
V54X, in particular V54I,N,W,Y,F,L;
G57S,A,V,L,I,F,Y,T;
G107X, in particular G107A,V,S,T,I,L,C;
G108X, in particular G108A,V,S,T,I,L;
A111V,I,L;
S168Y;
M197X, in particular Y,F,L,I,T,A,G.

15

4. The variant of any of claims 1-3, comprises the following mutations corresponding to at least one of the following mutations in the amino acid sequence shown in SEQ ID NO: 4:
T49X+A52X+V54N/I/L/Y/F/W+G107A.

20

5. The variant of claims 1-4, further comprising G108A.

25

6. The variant of claim 1-5, comprises the following mutations corresponding to at least one of the following mutations in the amino acid sequence shown in SEQ ID NO: 4:

T49L+G107A;
T49I+G107A;
T49L+G107A+V54I;
T49I+G107A+V54I;
A52S+V54N+T49L+G107A;
A52S+V54I+T49L+G107A;
A52S+T49L+G107A;
A52T+T49L+G107A;
A52S+V54N+T49I+G107A;
A52S+V54I+T49I+G107A;
A52S+T49I+G107A;
T49L+G108A;

T49I+G108A;
T49L+G108A+V54I;
T49I+G108A+V54I.

5 7. A variant of any of claims 1-6, wherein said variant has a
reduced capability of cleaving an oligo-saccharide substrate
close to the branching point as compared to the parent alpha-
amylase.

10 8. A variant of any of claims 1-7, which further exhibits
improved substrate specificity and/or improved specific activity
relative to the parent Termamyl-like alpha-amylase.

15 9. A variant of any of claims 1-8, wherein the parent alpha-
amylase is a hybrid alpha-amylase of SEQ ID NO: 4 and SEQ ID NO:
6.

20 10. The variant of any of claims 1-9, wherein the parent hybrid
alpha-amylase is a hybrid alpha-amylase comprising the 445 C-
terminal amino acid residues of the *B. licheniformis* alpha-
amylase shown in SEQ ID NO: 4 and the 37 N-terminal amino acid
residues of the alpha-amylase derived from *B. amyloliquefaciens*
shown in SEQ ID NO: 6.

25 11. The variant of any of claims 1-10, wherein the parent hybrid
Termamyl-like alpha-amylase further has the following
mutations: H156Y+A181T+N190F+A209V+Q264S (using the numbering
in SEQ ID NO: 4) or LE174.

30 12. The variant of any of claims 1-11, wherein the parent hybrid
Termamyl-like alpha-amylase further has the following
mutations: H156Y+A181T+N190F+A209V+Q264S+I201F (using the
numbering of SEQ ID NO: 4) or LS429.

35 13. A DNA construct comprising a DNA sequence encoding an alpha-
amylase variant according to any of claims 1-12.

14. A recombinant expression vector which carries a DNA construct according to claim 13.

15. A cell which is transformed with a DNA construct according to claim 13 or a vector according to claim 14.

16. A cell of claim 9, which is a microorganism, in particular a bacterium or a fungus, such as a gram positive bacterium such as *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus lentus*, *Bacillus brevis*, *Bacillus stearothermophilus*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus coagulans*, *Bacillus circulans*, *Bacillus lautus* or *Bacillus thuringiensis*.

17. A composition comprising:

(i) a mixture of the alpha-amylase from *B. licheniformis* having the sequence shown in SEQ ID NO: 4 with one or more variants of claims 1-12 derived from (as the parent Termamyl-like alpha-amylase) the *B. stearothermophilus* alpha-amylase having the sequence shown in SEQ ID NO: 8; or

(ii) a mixture of the alpha-amylase from *B. stearothermophilus* having the sequence shown in SEQ ID NO: 8 with one or more variants of claims 1-12 derived from one or more other parent Termamyl-like alpha-amylases; or

(iii) a mixture of one or more variants of claim 1-12 derived from (as the parent Termamyl-like alpha-amylase) the *B. stearothermophilus* alpha-amylase having the sequence shown in SEQ ID NO: 8 with one or more variants according to the invention derived from one or more other parent Termamyl-like alpha-amylases.

18. A composition comprising:

a mixture of one or more variants of claims 1-12 derived from (as the parent Termamyl-like alpha-amylase) the *B. stearothermophilus* alpha-amylase having the sequence shown in SEQ ID NO: 8 and a Termamyl-like alpha-amylase derived from the *B. licheniformis* alpha-amylase having the sequence shown in SEQ ID NO: 4.

19. A composition comprising:

a mixture of one or more variants of claims 1-12 derived from (as the parent Termamyl-like alpha-amylase) the *B. stearothermophilus* alpha-amylase having the sequence shown in SEQ ID NO: 8 and a hybrid alpha-amylase comprising a part of the *B. amyloliquefaciens* alpha-amylase shown in SEQ ID NO: 6 and a part of the *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4.

20. A composition comprising:

a mixture of one or more variants of claims 1-12 derived from (as the parent Termamyl-like alpha-amylase) a hybrid alpha-amylase comprising a part of the *B. amyloliquefaciens* alpha-amylase shown in SEQ ID NO: 6 and a part of the *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4.

21. A composition of claim 20, wherein the hybrid alpha-amylase is a hybrid alpha-amylase comprising the 445 C-terminal amino acid residues of the *B. licheniformis* alpha-amylase shown in SEQ ID NO: 4 and the 37 N-terminal amino acid residues of the alpha-amylase derived from *B. amyloliquefaciens* shown in SEQ ID NO: 6.

22. A composition of claim 21, wherein the hybrid alpha-amylase further has the following mutations:

H156Y+A181T+N190F+A209V+Q264S (using the numbering in SEQ ID NO: 4) or LE174.

23. A composition of claim 21, wherein the hybrid alpha-amylase further has the following mutations:

H156Y+A181T+N190F+A209V+Q264S+I201F as shown in SEQ ID NO: 2 or LE429.

24. A method for generating a variant of a parent Termamyl-like alpha-amylase, which variant exhibits a reduced capability of cleaving a substrate close to the branching point, and further exhibits improved substrate specificity and/or improved

specific activity relative to the parent, the method comprising:

- (a) subjecting a DNA sequence encoding the parent Termamyl-like alpha-amylase to random mutagenesis,
- (b) expressing the mutated DNA sequence obtained in step (a) in a host cell, and
- (c) screening for host cells expressing a mutated alpha-amylase which has increased stability at low pH and low calcium concentration relative to the parent alpha-amylase.

10

25. Use of an alpha-amylase variant of any of claims 1-12 or a composition of any of claims 17-23 for starch liquefaction; in detergent composition, such as laundry, dish washing and hard surface cleaning compositions; ethanol production, such as fuel, drinking and industrial ethanol production; desizing of textiles, fabrics or garments.

15

SEQUENCE LISTING

5 <110> Novo Nordisk A/S
 <120>
 <130>
 10 <160> 40
 <170> ParentIn Ver. 2.1
 <210> 1
 15 <211> 1443
 <212> DNA
 <213> *Bacillus amyloliquefaciens*
 <220>
 20 <221> CDS
 <222> (1)...(1443)
 <400> 1
 25 gta aat ggc acg ctg aag cag tat ttt gaa tgg tat acg ccg aac gac 46
 Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro Asn Asp
 1 5 10 15
 ggc cag cat tgg aaa cga ttt cag aat gat gcg gaa cat ttt tgg gat 96
 Gly Cln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu Ser Asp
 20 25 30
 30 atc ggt att act gcc gtc tgg att ccc ccc gca tat aag gga acg aac 144
 Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser
 35 40 45
 35 caa ggc gat gtg ggc tac egt qct tac gac ctt tat gat ttt ggg gag 192
 Gln Ala Asp Val Gly Tyr Ala Tyr Asp Leu Tyr Asp Leu Gly Glu
 50 55 60
 40 ttg cat caa aaa ggg acg gtt cgg aca aag tcc ggc aca aca gga gag 240
 Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu
 65 70 75 80
 45 ctg caa tct gmg atc aca agt ctt cat tcc cgc gac att aac gtt tcc 288
 Leu Cln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr
 85 90 95
 50 ggg gat gtg gtc atc aac cac aca aca ggc ggc gat gat gcg acc gaa gat 336
 Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp
 100 105 110
 55 gta acc gcg gtt gaa gtc gat ccc gct gac cgc aac cgc gta att tca 384
 Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser
 115 120 125
 60 gca gca ccc cta att aca gcc tgg aca cat ttt cat ttt ccg ggg ccc 432
 Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg
 130 135 140
 65 ggc agc aca tac ago gat ttt aag tgg tat tgg tec cat ttt gac ggs 480
 Gly Ser Thr Tyr Ser Asp Phe Lys Trp Tyr Trp Tyr His Phe Asp Gly
 145 150 160
 70 acc gat tgg gac gag tcc cgs aag ctg aac cgc att tat aag ttt cca 528
 Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln
 165 170 175

5	ggg aag aca tgg gat tgg gaa gtt tcc aac gaa trc ggc aac tat gat Gly Lys Thr Trp Asp Trp Glu Val Ser Asn Glu Phe Gly Asn Tyr Asp 186 185 190	576
10	tat tgg atg tat gcc gac ctt gat tat gag cat cct gat gtc gta gca Tyr Leu Met Tyr Ala Asp Phe Asp Tyr Asp His Pro Asp Val Val Ala 195 200 205	624
15	gag att aag aca tgg ggc act tgg tat gcc aat gac ctg caa ttg gac Glu Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Cln Leu Asp 210 215 220	672
20	act ttc cgt ctt gat gct gtc aas cac att aac ttt tcc ctt ttg cgg Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg 225 230 235 240	720
25	gat tgg gtt aat cat gtc aag gaa aac aag ggg aag gaa atg ttt acg Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr 245 250 255	768
30	gta gct gag tcc tgg tcg aat gad ttg ggc gcg ctg gaa aac taa tgg Val Ala Glu Trp Ser Asn Asp Ieu Gly Ala Leu Glu Asn Tyr Leu 260 265 270	816
35	aac aca aca aat ttc aac cat tca gtc ttt gad gtc cog ctt cat tat Asn Lys Thr Asn Phe Asn His Ser Val Asp Val Pro Leu His Tyr 275 280 285	854
40	cag ttc cat gca gca tcc aca cag gga ggc ggc tat gat atg agg aaa Gln Phe His Ala Ala Ser Thr Gln Gly Gly Tyr Asp Met Arg Lys 290 295 300	912
45	ttg ctc aac ggt acg gca gta tcc sag cat ccc tcc ssa tcc gtt sca Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser Val Thr 305 310 315 320	960
50	ttt gtc gat aac cat gat aca cay ccc ggg caa tcc ctt gag tcc act Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr 325 330 335	1008
55	gtc caa aca tgg ttt aag ccc ctt gct tac gct gtt ttt aat ccc sca agg Val Gln Thr Trp Phe Iys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg 340 345 350	1056
60	gaa tcc gga tac ccc cag gtt ttc tac ggg gat atg tcc ggg acg aca Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Ilys 355 360 365	1104
65	ggc gac tcc ccc ccc gaa att ccc gca ttg aac ccc aca att gaa ccc Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro 370 375 380	1152
70	atc tta aac gcc aga aca cag tat gcc tcc gca gca ccc cat gat aat Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr 385 390 395 400	1200
75	ccc gac ccc cat gac att gtc ggc tcc aca aag gaa ggc gac acg tcc Phe Asp His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser 405 410 415	1248
80	gtt gca aat tca ggt ttg gcc gca tta ata sca gac gga ccc ggt ggg Val Ala Asn Ser Gly Leu Ala Leu Ile Thr Asp Gly Pro Gly Gly 420 425 430	1296
85	gca aag cca atg tat gtc ggc egg cca aac gcc ggt gac aca tgg cat Ala Lys Arg Met Tyr Val Gly Arg Gla Asn Ala Gly Glu Thr Trp His 435 440 445	1344

5	gac att acc gga aac cgt tgg ggg cog gtt gtc atc att tcg gaa ggc Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly 450 455 460		1292
6	tgg gga gag ccc cac gta aac ggc ggg tgg gtt tca att tat gtt caa Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln 465 470 475 480		1449
10	aga Arg		1443
15	<210> 2 <211> 481 <212> PRT <213> <i>Bacillus amyloliquefaciens</i>		
20	<400> 2 Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro Asn Asp 1 5 10 15		
25	Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu Ser Asp 30 35 40 45		
30	Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser 35 40 45		
35	Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Gln 50 55 60		
40	Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu 65 70 75 80		
45	Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr 85 90 95		
50	Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp 100 105 110		
55	Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser 115 120 125		
60	Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg 130 135 140		
65	Gly Ser Thr Tyr Ser Asp Phe Lys Trp Tyr Trp Tyr His Phe Asp Gly 145 150 155 160		
70	Thr Asp Trp Asp Gln Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln 165 170 175		
75	Gly Lys Thr Trp Asp Trp Glu Val Ser Asn Glu Phe Gly Asn Tyr Asp 185 185 190 195		
80	Tyr Leu Met Tyr Ala Asp Phe Asp Tyr Asp His Pro Asp Val Val Ala 195 200 205		
85	Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp 210 215 220		
90	Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg 225 230 235 240		
95	Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr 245 250 255		

5

<210> 3
<211> 1920
50 <212> DNA
<213> *Bacillus licheniformis*

SS <220>
 <221> CGG
 <222> (421) ..(1872)

<400> 3
ccggaaatgtg ggaaatccaaa aataagcaaa agattgrcaa tcatgtcatg agccatgcgg 60
60 gagazggaaa aatcgtaata atgcaacgata ttatgcac ac gttcgcagat yctgctgaag 120
ggatattcaaa aaagctgaaa gcaaaaaggcc atccatgttgtt aactgtatct cagcttgaaag 180
65 aagtgtaaatggaa qcagagagggc tattgtatca atgagtagaa gcgccatatac ggcgcgttttc 240
ttttagccqaa aatataatggg aaaaatgttc ttgttaaaaat ttccggatata ttatacaacca 300

tcatatgttt vacatggaaa gggggaggaga accatgaaac aacaaaaacy gctttacgc 360
 cgtatgtgtga cgctgttatt tgccgcrcatc trctttgtgc ctcatttgtc agcagcgccg 420
 6 gca aat ctg aat ggg acy ctg atg cag tat ttt gaa tgg tac atg ccc 468
 Ala Asn Leu Asn Gly Thr Leu Met Gln Tyr Phe Glu Pro Tyr Met Pro
 1 5 10 15
 10 aat gac ggc caa cat tgg agg cgt ttg caa aac gac tgg gca tat ttg 516
 Asn Asp Gly Gln His Trp Arg Arg Leu Gln Asn Asp Ser Ala Tyr Leu
 20 25 30
 12 gct gaa cac ggt att acc gcc gtc tgg att ccc ccg gca tat aag gga 564
 Ala Glu His Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly
 35 40 45
 15 acc aac cca gcg gat gtc ggc tac ggt gct tac gac ctt tac gat tta 612
 Thr Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu
 50 55 60
 20 25 30
 18 ggg gag ttt cat caa aas ggg acy gtt cgg aca aag tac ggc aca aaa 660
 Gly Glu Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys
 65 70 75 80
 25 30 35
 22 gga gag ctg cca tct gcg atc aca agt ctt cat tcc cgc gac att aac 708
 Gly Glu Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn
 85 90 95
 30 35 40
 28 gtt tac ggg gat gtc gtc arc aac cac aca gga ggc gtc gat gat gcg acc 756
 Val Tyr Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr
 100 105 110
 35 40 45
 32 gaa gat gta acc gca acc gaa gtc gat ccc gct gac cgc aac cgg gta 804
 Glu Asp Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val
 115 120 125
 40 45 50
 36 art tca gga gaa cac cta art aca gcc tgg aca cat tcc cat ttt ccc 852
 Ile Ser Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Pro
 130 135 140
 45 50 55
 40 45 50
 38 ggg cgc ggc agc aca tac agc gat ttt aca tgg cat tgg tac cat ttt 900
 Gly Arg Gly Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe
 145 150 155 160
 50 55 60
 42 47 52
 44 gac gga acc gat tgg gac gag tcc cga aac ctg aac cgc atc bat aac 946
 Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys
 165 170 175
 55 60 65
 46 51 56
 48 53 58
 50 55 60
 52 57 62
 54 59 64
 56 61 66
 58 63 68
 60 65 70
 62 67 72
 64 69 74
 66 71 76
 68 73 78
 70 75 80
 72 77 82
 74 79 84
 76 81 86
 78 83 88
 80 85 90
 82 87 92
 84 89 94
 86 91 99
 88 93 100
 90 95 100
 92 97 100
 94 99 100
 96 101 100
 98 103 100
 100 105 100
 102 107 100
 104 109 100
 106 111 100
 108 113 100
 110 115 100
 112 117 100
 114 119 100
 116 121 100
 118 123 100
 120 125 100
 122 127 100
 124 129 100
 126 131 100
 128 133 100
 130 135 100
 132 137 100
 134 139 100
 136 141 100
 138 143 100
 140 145 100
 142 147 100
 144 149 100
 146 151 100
 148 153 100
 150 155 100
 152 157 100
 154 159 100
 156 161 100
 158 163 100
 160 165 100
 162 167 100
 164 169 100
 166 171 100
 168 173 100
 170 175 100
 172 177 100
 174 179 100
 176 181 100
 178 183 100
 180 185 100
 182 187 100
 184 189 100
 186 191 100
 188 193 100
 190 195 100
 192 197 100
 194 199 100
 196 201 100
 198 203 100
 200 205 100
 202 207 100
 204 209 100
 206 211 100
 208 213 100
 210 215 100
 212 217 100
 214 219 100
 216 221 100
 218 223 100
 220 225 100
 222 227 100
 224 229 100
 226 231 100
 228 233 100
 230 235 100
 232 237 100
 234 239 100
 236 241 100
 238 243 100
 240 245 100
 242 247 100
 244 249 100
 246 251 100
 248 253 100
 250 255 100
 252 257 100
 254 259 100
 256 261 100
 258 263 100
 260 265 100
 262 267 100
 264 269 100
 266 271 100
 268 273 100
 270 275 100
 272 277 100
 274 279 100
 276 281 100
 278 283 100
 280 285 100
 282 287 100
 284 289 100
 286 291 100
 288 293 100
 290 295 100
 292 297 100
 294 299 100
 296 301 100
 298 303 100
 300 305 100
 302 307 100
 304 309 100
 306 311 100
 308 313 100
 310 315 100
 312 317 100
 314 319 100
 316 321 100
 318 323 100
 320 325 100
 322 327 100
 324 329 100
 326 331 100
 328 333 100
 330 335 100
 332 337 100
 334 339 100
 336 341 100
 338 343 100
 340 345 100
 342 347 100
 344 349 100
 346 351 100
 348 353 100
 350 355 100
 352 357 100
 354 359 100
 356 361 100
 358 363 100
 360 365 100
 362 367 100
 364 369 100
 366 371 100
 368 373 100
 370 375 100
 372 377 100
 374 379 100
 376 381 100
 378 383 100
 380 385 100
 382 387 100
 384 389 100
 386 391 100
 388 393 100
 390 395 100
 392 397 100
 394 399 100
 396 401 100
 398 403 100
 400 405 100
 402 407 100
 404 409 100
 406 411 100
 408 413 100
 410 415 100
 412 417 100
 414 419 100
 416 421 100
 418 423 100
 420 425 100
 422 427 100
 424 429 100
 426 431 100
 428 433 100
 430 435 100
 432 437 100
 434 439 100
 436 441 100
 438 443 100
 440 445 100
 442 447 100
 444 449 100
 446 451 100
 448 453 100
 450 455 100
 452 457 100
 454 459 100
 456 461 100
 458 463 100
 460 465 100
 462 467 100
 464 469 100
 466 471 100
 468 473 100
 470 475 100
 472 477 100
 474 479 100
 476 481 100
 478 483 100
 480 485 100
 482 487 100
 484 489 100
 486 491 100
 488 493 100
 490 495 100
 492 497 100
 494 499 100
 496 501 100
 498 503 100
 500 505 100
 502 507 100
 504 509 100
 506 511 100
 508 513 100
 510 515 100
 512 517 100
 514 519 100
 516 521 100
 518 523 100
 520 525 100
 522 527 100
 524 529 100
 526 531 100
 528 533 100
 530 535 100
 532 537 100
 534 539 100
 536 541 100
 538 543 100
 540 545 100
 542 547 100
 544 549 100
 546 551 100
 548 553 100
 550 555 100
 552 557 100
 554 559 100
 556 561 100
 558 563 100
 560 565 100
 562 567 100
 564 569 100
 566 571 100
 568 573 100
 570 575 100
 572 577 100
 574 579 100
 576 581 100
 578 583 100
 580 585 100
 582 587 100
 584 589 100
 586 591 100
 588 593 100
 590 595 100
 592 597 100
 594 599 100
 596 601 100
 598 603 100
 600 605 100
 602 607 100
 604 609 100
 606 611 100
 608 613 100
 610 615 100
 612 617 100
 614 619 100
 616 621 100
 618 623 100
 620 625 100
 622 627 100
 624 629 100
 626 631 100
 628 633 100
 630 635 100
 632 637 100
 634 639 100
 636 641 100
 638 643 100
 640 645 100
 642 647 100
 644 649 100
 646 651 100
 648 653 100
 650 655 100
 652 657 100
 654 659 100
 656 661 100
 658 663 100
 660 665 100
 662 667 100
 664 669 100
 666 671 100
 668 673 100
 670 675 100
 672 677 100
 674 679 100
 676 681 100
 678 683 100
 680 685 100
 682 687 100
 684 689 100
 686 691 100
 688 693 100
 690 695 100
 692 697 100
 694 699 100
 696 701 100
 698 703 100
 700 705 100
 702 707 100
 704 709 100
 706 711 100
 708 713 100
 710 715 100
 712 717 100
 714 719 100
 716 721 100
 718 723 100
 720 725 100
 722 727 100
 724 729 100
 726 731 100
 728 733 100
 730 735 100
 732 737 100
 734 739 100
 736 741 100
 738 743 100
 740 745 100
 742 747 100
 744 749 100
 746 751 100
 748 753 100
 750 755 100
 752 757 100
 754 759 100
 756 761 100
 758 763 100
 760 765 100
 762 767 100
 764 769 100
 766 771 100
 768 773 100
 770 775 100
 772 777 100
 774 779 100
 776 781 100
 778 783 100
 780 785 100
 782 787 100
 784 789 100
 786 791 100
 788 793 100
 790 795 100
 792 797 100
 794 799 100
 796 801 100
 798 803 100
 800 805 100
 802 807 100
 804 809 100
 806 811 100
 808 813 100
 810 815 100
 812 817 100
 814 819 100
 816 821 100
 818 823 100
 820 825 100
 822 827 100
 824 829 100
 826 831 100
 828 833 100
 830 835 100
 832 837 100
 834 839 100
 836 841 100
 838 843 100
 840 845 100
 842 847 100
 844 849 100
 846 851 100
 848 853 100
 850 855 100
 852 857 100
 854 859 100
 856 861 100
 858 863 100
 860 865 100
 862 867 100
 864 869 100
 866 871 100
 868 873 100
 870 875 100
 872 877 100
 874 879 100
 876 881 100
 878 883 100
 880 885 100
 882 887 100
 884 889 100
 886 891 100
 888 893 100
 890 895 100
 892 897 100
 894 899 100
 896 901 100
 898 903 100
 900 905 100
 902 907 100
 904 909 100
 906 911 100
 908 913 100
 910 915 100
 912 917 100
 914 919 100
 916 921 100
 918 923 100
 920 925 100
 922 927 100
 924 929 100
 926 931 100
 928 933 100
 930 935 100
 932 937 100
 934 939 100
 936 941 100
 938 943 100
 940 945 100
 942 947 100
 944 949 100
 946 951 100
 948 953 100
 950 955 100
 952 957 100
 954 959 100
 956 961 100
 958 963 100
 960 965 100
 962 967 100
 964 969 100
 966 971 100
 968 973 100
 970 975 100
 972 977 100
 974 979 100
 976 981 100
 978 983 100
 980 985 100
 982 987 100
 984 989 100
 986 991 100
 988 993 100
 990 995 100
 992 997 100
 994 999 100
 996 1000 100
 998 1002 100
 1000 1004 100
 1002 1006 100
 1004 1008 100
 1006 1012 100
 1008 1016 100
 1010 1020 100
 1012 1024 100
 1014 1028 100
 1016 1032 100
 1018 1036 100
 1020 1040 100
 1022 1044 100
 1024 1048 100
 1026 1052 100
 1028 1056 100
 1030 1060 100
 1032 1064 100
 1034 1068 100
 1036 1072 100
 1038 1076 100
 1040 1080 100
 1042 1084 100
 1044 1088 100
 1046 1092 100
 1048 1096 100
 1050 1100 100
 1052 1104 100
 1054 1108 100
 1056 1112 100
 1058 1116 100
 1060 1120 100
 1062 1124 100
 1064 1128 100
 1066 1132 100
 1068 1136 100
 1070 1140 100
 1072 1144 100
 1074 1148 100
 1076 1152 100
 1078 1156 100
 1080 1160 100
 1082 1164 100
 1084 1168 100
 1086 1172 100
 1088 1176 100
 1090 1180 100
 1092 1184 100
 1094 1188 100
 1096 1192 100
 1098 1196 100
 1100 1200 100
 1102 1204 100
 1104 1208 100
 1106 1212 100
 1108 1216 100
 1110 1220 100
 1112 1224 100
 1114 1228 100
 1116 1232 100
 1118 1236 100
 1120 1240 100
 1122 1244 100
 1124 1248 100
 1126 1252 100
 1128 1256 100
 1130 1260 100
 1132 1264 100
 1134 1268 100
 1136 1272 100
 1138 1276 100
 1140 1280 100
 1142 1284 100
 1144 1288 100
 1146 1292 100
 1148 1296 100
 1150 1300 100
 1152 1304 100
 1154 1308 100
 1156 1312 100
 1158 1316 100
 1160 1320 100
 1162 1324 100
 1164 1328 100
 1166 1332 100
 1168 1336 100
 1170 1340 100
 1172 1344 100
 1174 1348 100
 1176 1352 100
 1178 1356 100
 1180 1360 100
 1182 1364 100
 1184 1368 100
 1186 1372 100
 1188 1376 100
 1190 1380 100
 1192 1384 100
 1194 1388 100
 1196 1392 100
 1198 1396 100
 1200 1400 100
 1202 1404 100
 1204 1408 100
 1206 1412 100
 1208 1416 100
 1210 1420 100
 1212 1424 100
 1214 1428 100
 1216 1432 100
 1218 1436 100
 1220 1440 100
 1222 1444 100
 1224 1448 100
 1226 1452 100
 1228 1456 100
 1230 1460 100
 1232 1464 100
 1234 1468 100
 1236 1472 100
 1238 1476 100
 1240 1480 100
 1242 1484 100
 1244 1488 100
 1246 1492 100
 1248 1496 100
 1250 1500 100
 1252 1504 100
 1254 1508 100
 1256 1512 100
 1258 1516 100
 1260 1520 100
 1262 1524 100
 1264 1528 100
 1266 1532 100
 1268 1536 100
 1270 1540 100
 1272 1544 100
 1274 1548 100
 1276 1552 100
 1278 1556 100
 1280 1560 100
 1282 1564 100
 1284 1568 100
 1286 1572 100
 1288 1576 100
 1290 1580 100
 1292 1584 100
 1294 1588 100
 1296 1592 100
 1298 1596 100
 1300 1600 100
 1302 1604 100
 1304 1608 100
 1306 1612 100
 1308 1616 100
 1310 1620 100
 1312 1624 100
 1314 1628 100
 1316 1632 100
 1318 1636 100
 1320 1640 100
 1322 1644 100
 1324 1648 100
 1326 1652 100
 1328 1656 100
 1330 1660 100
 1332 1664 100
 1334 1668 100
 1336 1672 100
 1338 1676 100
 1340 1680 100
 1342 1684 100
 1344 1688 100
 1346 1692 100
 1348 1696 100
 1350 1700 100
 1352 1704 100
 1354 1708 100
 1356 1712 100
 1358 1716 100
 1360 1720 100
 1362 1724 100
 1364 1728 100
 1366 1732 100
 1368 1736 100
 1370 1740 100
 1372 1744 100
 1374 1748 100
 1376 1752 100
 1378 1756 100
 1380 1760 100
 1382 1764 100
 1384 1768 100
 1386 1772 100
 1388 1776 100
 1390 1780 100
 1392 1784 100
 1394 1788 100
 1396 1792 100
 1398 1796 100
 1400 1800 100
 1402 1804 100
 1404 1808 100
 1406 1812 100
 1408 1816 100
 1410 1820 100
 1412 1824 100
 1414 1828 100
 1416 1832 100
 1418 1836 100
 1420 1840 100
 1422 1844 100
 1424 1848 100
 1426 1852 100
 1428 1856 100
 1430 1860 100
 1432 1864

	ttt aac gta gct gaa tat tgg cag aat gag ttt ggc ggc ctg gaa aac Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn 260 265 270	1236
5	cat ttt aac aca aac ptt aat cat tca gtc ttt gac gtc ccc ctt Tyr Leu Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu 275 280 285	1284
10	cac tat cag trc cat gct gca tcg aca cag gga ggc ggc bat gat atg His Tyr Gln Phe His Ala Ala Ser Thr Gln Gly Gly Tyr Asp Met 290 295 300	1332
15	agg aaa ttt gtc aac ggt agg gtc qtt tcc aag cat ccc ttt aca tcc Arg Lys Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser 305 310 315 320	1380
20	gtt aca ttt gtc gat aac cat gat aca ccc ccc ggg ccc tcg ctt ggg Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu 325 330 335	1428
25	tcc act gtc ccc aca tgg ttt aag ccc ctt gct tac got ttt att ctc Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu 340 345 350	1476
30	aca aca gaa tcc gga tac ccc ccc gtt tcc tac ggg gat atg tac ggg Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly 355 360 365	1524
35	acg aca gga gac tcc ccc ccc acc ccc gcc ttg aca cac aca att Thr Lys Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile 370 375 380	1572
40	gaa ccg atc ttt aca gcg aga aca ccc tat gcc tac gga gca ccc cat Glu Pro Ile Ile Lys Ala Arg Lys Glu Tyr Ala Tyr Gly Ala Glu His 385 390 395 400	1620
45	gat tat tcc gac ccc cat gac att gtc ggc tgg aca agg gaa ggc gac Asp Tyr Phe Asp His Asp Ile Val Gly Trp Thr Arg Glu Gly Asp 405 410 415	1668
50	agg tcc gtt gca sat tcc ggt ttt gcc gca ttt atc aca gac gga ccc Ser Ser Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro 420 425 430	1716
55	gtt ggg gca aac ccc atg tat gtc ggc ccc aac gcc ggc gag aca Gly Gly Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr 435 440 445	1764
60	tgg cat gac att acc gga aac cgt tcg gag ccc gtt gtc atc aat tcc Trp His Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser 450 455 460	1812
65	gaa ggc tgg gga gag ttt ccc gca aac ggc ggg tcc gtt tca att tcc Glu Glu Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr 465 470 475 480	1860
70	gtt aca aga tag aac ggc aac ggg agg aac ggg att ccc gtt ttt Val Glu Arg	1812
75	ttt atccc	1820
80	<210> 4 <211> 483 <212> PRT <213> <i>Bacillus licheniformis</i>	

<400> 4
 Ala Asn Leu Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro
 1 5 10 15
 Asn Asp Gly Gln His Trp Arg Arg Leu Gln Asn Asp Ser Ala Tyr Leu
 20 25 30
 Ala Gln His Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly
 35 40 45
 Thr Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu
 50 55 60
 Gly Gln Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys
 65 70 75 80
 Gly Gln Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn
 85 90 95
 Val Tyr Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr
 100 105 110
 Glu Asp Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val
 115 120 125
 Ile Ser Gly Glu His Leu Ile Lys Ala Trp Thr His Phe His Phe Pro
 130 135 140
 Gly Arg Gly Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe
 145 150 155 160
 Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys
 165 170 175
 Phe Gln Gly Lys Ala Trp Asp Trp Glu Val Ser Asn Gln Asn Gly Asn
 180 185 190
 Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val
 195 200 205
 Ala Ala Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln
 210 215 220
 Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe
 225 230 235 240
 Leu Arg Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Gln Met
 245 250 255 260
 Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn
 265 270 275
 Tyr Leu Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu
 275 280 285
 His Tyr Gln Phe His Ala Ala Ser Thr Gln Gly Gly Tyr Asp Met
 290 295 300
 Arg Lys Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ser
 305 310 315 320
 Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu
 325 330 335
 Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu
 340 345 350

	cgt acc ctt sag ttt cgt ggg gas gga aaa ggc tgg tat tgg gaa gta Arg Ile Phe Lys Arg Cys Glu Gly Lys Ala Trp Asp Trp Glu Val 175 180 185	1470
3	tcc agt gaa aac ggc aac tat gac tat tta atg tat gct gat gtt gac Ser Ser Glu Asn Cys Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp 190 195 200	1488
10	tac gag cac cct gat gtc gtc gca gag aca aaa aaa tgg ggt aac tgg Tyr Asp His Pro Asp Val Val Ala Glu Thr Lys Trp Cys Ile Trp 205 210 215	1516
18	tat gcg aat gaa ctg tca ttg gac ggc ttc cgt att gat gcc gcc aac Tyr Ala Asn Glu Leu Ser Leu Asp Gly Phe Arg Ile Asp Ala Ala Lys 220 225 230	1564
26	cac att aac ttt tca ttt ctg cgt gat tgg gtt cag ggc gac aga cag His Ile Lys Phe Ser Phe Leu Arg Asp Trp Val Gln Ala Val Arg Gln 235 240 245 250	1612
	ggc acc gga aaa gaa atg ttt acg gtc ggg gag tac tgg cag aat aat Ala Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Gln Asn Asn 255 260 265	1660
28	gcc egg aca ccc gaa aac tac ttg aat aac aca agc ttt aat caa tcc Ala Gly Lys Leu Asn Tyr Leu Asn Lys Thr Ser Phe Asn Gln Ser 270 275 280	1708
36	gtg ctt gat gtt ccg ctt ctc aat tta cag ggg gct tcc tcc caa Val Phe Val Pro Leu His Phe Asn Leu Gln Ala Ala Ser Ser Gln 285 290 295	1756
38	gga ggc gpa tac gat atg agg cgt ttg ctg gac ggt acc gtt gtc tcc Gly Gly Gly Tyr Asp Met Arg Arg Leu Asp Gly Thr Val Val Ser 300 305 310	1804
40	agg cat ccg gaa aag ggc gtt ecc ttg gtt gaa aat cat gac ecc cag Arg His Pro Glu Lys Ala Val Thr Phe Val Glu Asn His Asp Thr Gln 315 320 325 330	1852
	cgg gga cag tca ttg gaa tgg aca gtc cca act tgg ttt aaa ccc ott Pro Gly Gln Ser Leu Glu Ser Thr Val Gln Thr Trp Phe Lys Pro Leu 335 340 345	1900
46	gca tac gcc ctt att ttg aca aga gaa tcc ggt tat cct cag gtg ttc Ala Tyr Ala Phe Ile Leu Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe 350 355 360	1948
50	tat ggg gat atg tac ggg aca aca ggg aca ttg cca aag gca aac ccc Tyr Cys Asp Met Tyr Cys Thr Lys Cys Thr Ser Pro Lys Glu Ile Pro 365 370 375	1986
58	tca ctc aca gat aat ata gag ccg att tta aca ggg cgt aag gag tac Ser Leu Lys Asp Asn Ile Glu Pro Ile Leu Lys Ala Arg Lys Glu Tyr 380 385 390	2044
60	gca tac ggg ccc cag ccc gat tat att gac ccc ccg gat gtc aca gga Ala Tyr Cys Pro Gln His Asp Tyr Ile Asp His Pro Asp Val Ile Cys 395 400 405 410	2092
	tgg acc agg gaa ggt gac agc tcc gcc gcc aas tca ggt ttg gcc gct Trp Thr Arg Glu Asp Ser Ser Ala Ala Lys Ser Gly Leu Ala Ala 415 420 425	2140
68	ttt aca acc gac gga ccc ggc gga tca aag cgg atg tat gcc ggc ctg Leu Ile Thr Asp Gly Pro Gly Ser Lys Arg Met Tyr Ala Cys Leu	2188

	430	435	440	
8	aaa bat gcc ggc gag sca tgg tat gac ata acg ggc aac cgt tca gat Lys Asn Ala Gly Gln Thr Trp Tyr Asp Ile Thr Gly Asn Arg Ser Asp 445 450 455			2236
16	act gta aac atc gga tct gac ggc tgg gga gag ttt cat gta aac gat Thr Val Lys Ile Gly Ser Asp Gly Trp Gly Glu His Val Asn Asp 460 465 470			2284
18	ggg tcc gtc tcc att tat gtt cag aac taa ggtaataaaaa aaadacccccc Gly Ser Val Ser Ile Tyr Val Glu Lys 475 480			2334
19	aagctgagtc cgggtatcag cttggagggtg cgtttatccc ttccggccgtt tgacaaggc 2394 ggcatcaggt gtgacaaata cgggtatgcgt gctgtccatag gtgacaaatc cggggatc 2454 gcggtttggc ttttcacat gtttgttttt tgtaaatca acaggcacgg agccggatc 2514 20 tttcgccatcg gaaaaataaag cggcgatcgt agctgtttcc satatggatt gttcatcggy 2574 atcgatgtt ttaatcacaat cgtggggatcc 3604			
26	<210> 6 <211> 483 <212> PRT <213> Bacillus amyloliquefaciens			
30	<400> 6 Val Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Thr Pro Asn Asp 1 5 10 15			
35	Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ala Glu His Leu Ser Asp 20 25 30			
40	Ile Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Leu Ser 35 40 45			
45	Gln Ser Asp Asn Gly Tyr Gly Pro Tyr Asp Leu Tyr Asp Leu Gly Glu 50 55 60			
50	Phe Gln Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ser Glu 65 70 75 80			
55	Leu Gln Asp Ala Ile Gly Ser Leu His Ser Arg Asn Val Gln Val Tyr 85 90 95			
60	Gly Asp Val Val Leu Asn His Lys Ala Gly Ala Asp Ala Thr Glu Asp 100 105 110			
65	Val Thr Ala Val Glu Val Asn Pro Ala Asn Arg Asn Gln Thr Ser 115 120 125			
70	Gln Gln Tyr Gln Ile Lys Ala Trp Thr Asp Phe Arg Phe Pro Gly Arg 130 135 140			
75	Gly Asn Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly 145 150 155 160			
80	Ala Asp Trp Asp Glu Ser Arg Lys Ile Ser Arg Ile Phe Lys Phe Arg 165 170 175			
85	Gly Gln Gly Lys Ala Trp Asp Trp Glu Val Ser Ser Glu Asn Gly Asn 180 185 190			

Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Tyr Asp His Pro Asp Val
 195 200 205
 Val Ala Glu Thr Lys Lys Trp Gly Ile Trp Tyr Ala Asn Glu Leu Ser
 210 215 220
 Leu Asp Gly Phe Arg Ile Asp Ala Ala Lys His Ile Lys Phe Ser Phe
 225 230 235 240
 10 Leu Arg Asp Trp Val Gln Ala Val Arg Gln Ala Thr Gly Lys Glu Met
 245 250 255
 Phe Thr Val Ala Glu Tyr Trp Gln Asn Asn Ala Gly Lys Leu Glu Asn
 260 265 270
 15 Tyr Leu Asn Lys Thr Ser Phe Asn Gln Ser Val Phe Asp Val Pro Leu
 275 280 285
 His Phe Asn Leu Gln Ala Ala Ser Ser Gln Gly Gly Tyr Asp Met
 290 295 300
 Arg Arg Leu Leu Asp Gly Thr Val Val Ser Arg His Pro Glu Lys Ala
 305 310 315 320
 25 Val Thr Phe Val Glu Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu
 325 330 335
 Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu
 340 345 350
 30 Thr Arg Gln Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly
 355 360 365
 Thr Lys Gly Thr Ser Pro Lys Glu Ile Pro Ser Leu Lys Asp Asn Ile
 370 375 380
 Glu Pro Ile Leu Lys Ala Arg Lys Glu Tyr Ala Tyr Gly Pro Gln His
 385 390 395 400
 40 Asp Tyr Ile Asp His Pro Asp Val Ile Gly Trp Thr Arg Gly Gly Asp
 405 410 415
 Ser Ser Ala Ala Lys Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro
 420 425 430
 45 Gly Gly Ser Lys Arg Met Tyr Ala Gly Leu Lys Asn Ala Gly Glu Thr
 435 440 445
 Tryp Tyr Asp Ile Thr Gly Asn Arg Ser Asp Thr Val Lys Ile Gly Ser
 450 455 460
 Asp Gly Trp Gly Gln Phe His Val Asn Asp Gly Ser Val Ser Ile Tyr
 465 470 475 480
 50 Val Gln Lys

<210> 7
 <211> 1548
 <212> DNA
 <213> *Bacillus stearothermophilus*
 <220>
 <221> CDS
 <222> (1)...(1548)

<400> 7
 gcc gca ccc ttt acc ggc acc atg atg cag tat ttt gaa tgg tac ttg 48
 Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu
 1 S 10 15
 5
 ccg gat gat ggc acc tta tgg acc aaa gtg gcc aat gaa gcc aac aac 96
 Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn
 26 25 30
 10 tta tcc agc ctt aac acc acc gct ctt tgg ctg ccc gct tac aaa 144
 Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys
 35 40 45
 15 gga aca aac cgc aac gta ggg tac gga gta tac gac ttt tat gac 192
 Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp
 50 55 60
 20 ccc ggc gaa ttc aat caa aac ggg acc gtc cgg aca aac tac gga aca 240
 Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr
 65 70 75 80
 25 aac gct cca cat ott cca gcc att cca gcc gcc cac gct ggt gga atg 288
 Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met
 85 90 95
 30 acg gaa tgg tgg gac gtc gug ttc gac cat aac ggc ggc gct gac ggc 336
 Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly
 100 105 110
 35 acg gaa tgg tgg gac gtc gug ttc gac cat aac ggc ggc gct gac ggc 384
 Thr Glu Trp Val Asp Ala Val Gln Val Asn Pro Ser Asp Arg Asn Gln
 115 120 125
 40 gaa atc ccc ggc acc cat cca ccc gca tgg acg aaa ttt gar ttt 432
 Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe
 130 135 140
 45 ccc ggg ccc ggc acc acc tac tcc aco ttt aac tgg ccc tgg tac cat 480
 Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His
 145 150 155 160
 50 ttt gag ggc gtt gat tgg gac gaa aco cya aac tgg aac cgg acc tac 528
 Phe Asp Gly Val Asp Tep Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr
 165 170 175
 55 aac tcc ccc ggc acc ggc aac gac tgg gat tgg gaa gta gac aco gaa 576
 Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu
 180 185 190
 60 aac ggg gac cat gac tac tta atg bat gcc gac ott gat atg gat cat 624
 Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His
 195 200 205
 65 ccc gaa gtc gtc acc gag ctg aac aac tgg ggg aac tgg bat gta aac 672
 Pro Glu Val Val Thr Gln Leu Lys Asn Trp Gly Lys Trp Tyr Val Asn
 210 215 220
 70 aca aco aac bat gat ggg ttc cgg ott gat gca gtc aac cat att aac 720
 Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys
 225 230 235 240
 75 ttc aco aac bat cat gat tgg trg tgg tat gtg cgt tct cag att aac 768
 Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln Thr Gly
 245 250 255
 80 aac aco cta ttt acc gtc ggg gaa tat tgg aco bat gac aco aac aac 816
 Lys Pro Leu Phe Thr Val Gly Gln Tyr Trp Ser Tyr Asp Ile Asn Lys

	260	265	270	
5	tgc cac aat tac att acg aaa aca gac gga acg atg tct ttc tcc gar Leu His Asn Tyr Ile Thr Lys Thr Asp Gly Thr Met Ser Leu Phe Asp 375 280 285			864
10	gcc ccc ttc tac acc aas ccc tat socc gct tcc aaa tcc ggg gyc gca Ala Phe Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Ala 290 295 300			912
15	tcc gat atg cgc acg ttc atg acc aat act ctc atg aaa gat caa ccc Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro 305 310 315 320			960
20	acc ttg gac gtc acc ttc gtc gat aat cat gag acc gaa ccc ggc cca Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln 325 330 335			1008
25	ggc cgg cag tcc tgg gtc gac cca tgg ttc aaa ccc ttg gct tac gcc Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala 340 345 350			1056
30	ttt att cca act cgg cgg gaa gga tac cgg tgg ttc tti tac ggt gag Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp 355 360 365			1104
35	tat tat ggc att cca cca tat sac att cct tgg ctg aaa agc aaa atc Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile 370 375 380			1152
40	gtt ccc ctc ctc atc ggg cgg agg gag tat gct tac gga acg cca cat Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Glu Thr Gln His 385 390 395 400			1200
45	gat tat ctt gag ccc tcc gac atc atc ggg tgg aca agg gaa ggg ggc Asp Tyr Leu Asp His Ser Asp Ile Ile Glu Trp Thr Arg Glu Gly Gly 405 410 415			1248
50	act gaa aac cca gga tcc gga ctg gcc gca ctg atc acc gat ggg ccc Thr Glu Lys Pro Gly Ser Gly Ieu Ala Ala Leu Ile Thr Asp Gly Pro 420 425 430			1296
55	ggg gga aac tgg aag tac gtt gtc aac cca aac gct gga aac gtc Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val 435 440 445			1344
60	tcc tat gag ctt acc ggc aac cgg aat gac acc gtc acc atc acc aac Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser 450 455 460			1392
65	gat gga tgg ggg gaa tcc aac gtc aat ggc ggt tgg gtc tgg gtt tgg Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp 465 470 475 480			1440
70	gtt acc aca aac acg acc gtt tct acc aat gat agg ccc atc acc aca acc Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Arg Pro Ile Thr Thr 485 490 495			1488
75	cca ccc tgg act ggt gaa tcc gtc cgg tgg acc gaa cca cgg ttg gtc Arg Pro Trp Thr Gly Glu Phe Val Arg Trp Thr Glu Pro Arg Leu Val 500 505 510			1536
80	gca tgg act tga Ala Trp Pro 515			1548

<210> 8
 <211> 516
 <212> PFT
 <213> *Bacillus stearothermophilus*
 9
 <400> 8
 Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu
 1 5 10 15
 10 Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn
 20 25 30
 Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys
 35 40 45
 15 Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp
 50 55 60
 Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr
 20 65 70 75 80
 Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met
 85 90 95
 25 Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly
 100 105 110
 Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln
 115 120 125
 30 Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe
 130 135 140
 Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His
 35 145 150 155 160
 Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr
 165 170 175
 40 Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu
 180 185 190
 Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His
 195 200 205
 45 Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr Val Asn
 210 215 220
 Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys
 50 225 230 235 240
 Phe Ser Phe Phe Asp Trp Leu Ser Tyr Val Arg Ser Gln Thr Gly
 245 250 255
 55 Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile Asp Lys
 260 265 270
 Leu His Asn Tyr Ile Thr Lys Thr Asp Gly Thr Met Ser Leu Phe Asp
 275 280 285
 60 Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Ala
 290 295 300
 Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro
 65 305 310 315 320
 Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln

	329	330	335
5	Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala 340 345 350		
10	Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp 355 360 365		
15	Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile 370 375 380		
20	Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His 385 390 395 400		
25	Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Gly 405 410 415		
30	Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro 420 425 430		
35	Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val 435 440 445		
40	Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser 450 455 460		
45	Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp 465 470 475 480		
50	Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Arg Pro Ile Thr Thr 485 490 495		
55	Arg Pro Trp Thr Gly Glu Phe Val Arg Trp Thr Gln Pro Arg Leu Val 500 505 510		
60	Ala Trp Pro 515		

```
40      <210> 9
        <211> 31
        <212> DNA
        <213> Artificial Sequence
45      <220>
        <223> Description of Artificial Sequence: Primer
                <400> 9
50      ggccgcgtggcc acccgtagcccc caatcccgctt g
```

```
55 <210> 10
      <211> 36
      <212> DNA
      <213> Artificial Sequence

      <220>
60 <223> Description of Artificial Sequence: Primer
      <400> 10
      ggtagttggcc accgttgcggcc caatccccatt ggcttcgg 36
```

<212> DNA
<213> Artificial Sequence

6 <220>
<223> Description of Artificial Sequence: Primer

<400> 11
cgtcgactgg tgagtactca accaaagtc 28

10 <210> 12
<211> 31
<212> DNA
<213> Artificial Sequence

15 <220>
<223> Description of Artificial Sequence: Primer

<400> 12
ggtcgttaggc accgttagccc tcatccgctt g 31

20 <210> 13
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

30 <400> 13
ggtcgttaggc accgttagccc atatccgctt g 31

35 <210> 14
<211> 31
<212> DNA
<213> Artificial Sequence

40 <220>
<223> Description of Artificial Sequence: Primer

<400> 14
ggtcgttaggc accgttagccc atatccgctt g 31

45 <210> 15
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

50 <400> 15
gcagcgttggc accgttatacg aagaggccacg tcaaaaac 36

55 <210> 16
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 16

catacggtgcc gaatttcattt gaaaacrrccc 30

5 <210> 17
<211> 34
<212> DNA
<213> Artificial Sequence

10 <220>
<223> Description of Artificial Sequence: Primer

<400> 17
catacggtgcc gaatttcaggg gaaacttccc aatc 34

15 <210> 18
<211> 41
<212> DNA
<213> Artificial Sequence

20 <220>
<223> Description of Artificial Sequence: Primer

25 <400> 18
ccgcgcgcggc ggaaatcaaa ttttgtccag gcttttaattt g 41

30 <210> 19
<211> 32
<212> DNA
<213> Artificial Sequence

35 <220>
<223> Description of Artificial Sequence: Primer

<400> 19
caaaaatggta ccaataccac ttssseatcgc tg 33

40 <210> 20
<211> 29
<212> DNA
<213> Artificial Sequence

45 <220>
<223> Description of Artificial Sequence: Primer

<400> 20

cttccccaaatc ccaaggccatc ctttgaaac 29

50 <210> 21
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

60 <400> 21
ctttttttttt gttttttttt cttttttttt 36

65 <210> 22
<211> 38
<212> DNA
<213> Artificial Sequence

<203> Description of Artificial Sequence: Primer
400 <400> 33
gcaccaagccgaa taacggctac ggtgc 25
5
<210> 34
<211> 28
<212> DNA
10 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer
15 <400> 34
gaaacgagccca atcggacgtcg ggctacgg 26

<210> 35
20 <211> 32
<212> DNA
<213> Artificial Sequence

<220>
25 <223> Description of Artificial Sequence: Primer

<400> 35
ggaaacgagcc aatcggataaa cggctacggt gc 32

30 <210> 36
<211> 25
<212> DNA
<213> Artificial Sequence
35
<220>
40 <223> Description of Artificial Sequence: Primer

<400> 36
45 gcatataagg gactgagccca agcgg 36

<210> 37
<211> 25
45 <212> DNA
<213> Artificial Sequence

<220>
50 <223> Description of Artificial Sequence: Primer

<400> 37
55 caaccacaaaaa gccccccgtg atgcgg 25

<210> 38
<211> 41
55 <212> DNA
<213> Artificial Sequence

<220>
60 <223> Description of Artificial Sequence: Primer

<400> 38
65 gcatataagg gactgagccca atcggataaa cggctacggt c 41

<210> 39

<211> 28
<212> DNA
<213> Artificial Sequence

5 <220>
<223> Description of Artificial Sequence: Primer

<400> 39
gaacggagccg atcggacgcg ggctacgg

28

10

<210> 40
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

15

<400> 46
gaacggagccg aaacggacgtg ggctacgg

28

20

25